

# ERROR ESTIMATES FOR APPROXIMATE SOLUTIONS TO BELLMAN EQUATIONS ASSOCIATED WITH CONTROLLED JUMP-DIFFUSIONS

ESPEN R. JAKOBSEN, KENNETH H. KARLSEN, AND CLAUDIA LA CHIOMA

**ABSTRACT.** We derive error estimates for approximate (viscosity) solutions of Bellman equations associated to controlled jump-diffusion processes, which are fully nonlinear integro-partial differential equations. Two main results are obtained: (i) error bounds for a class of monotone approximation schemes, which includes finite difference schemes, and (ii) bounds on the error induced when the original Lévy measure is replaced by a finite measure with compact support, an approximation process that is commonly used when designing numerical schemes for integro-partial differential equations. In our proofs we use and extend techniques introduced by Krylov and Barles-Jakobsen.

## CONTENTS

1. Introduction	1
2. Optimal control of jump-diffusion processes	5
2.1. The optimal control problem	5
2.2. Lévy processes	6
3. Preliminaries	8
4. An abstract framework for obtaining error bounds	10
5. Finite difference schemes	14
5.1. Remarks	16
5.2. The Newton-Cotes formulas	17
5.3. Proof of Theorem 5.1	18
6. General Lévy measures	21
6.1. Treatment of the operators $\mathcal{J}_r^\theta$ and $\mathcal{J}^{\theta,R}$	22
6.2. Error estimates	23
7. Appendix: Some technical results.	24
References	30

## 1. INTRODUCTION

We are interested in deriving error estimates for numerical schemes for fully nonlinear degenerate elliptic integro-partial differential equation of Bellman type. These nonlocal equations take the form

$$(1.1) \quad H(x, u(x), Du(x), D^2u(x), u(\cdot)) = 0 \quad \text{in } \mathbb{R}^N,$$

---

*Date:* September 8, 2005.

*2000 Mathematics Subject Classification.* Primary: 45K05, 49L25, 65M12, 65L70.

*Key words and phrases.* Integro-partial differential equation, viscosity solution, numerical scheme, stability, error estimate, stochastic optimal control, Lévy process, Bellman equation.

This work is supported by the European network HYKE, contract HPRN-CT-2002-00282. The research of E. R. Jakobsen is supported by the Research Council of Norway through grant no 151608/432. The research of K. H. Karlsen is supported by an Outstanding Young Investigators Award from the Research Council of Norway. This work was done while C. La Chioma visited the Centre of Mathematics for Applications (CMA) at the University of Oslo, Norway.

where, for any  $(x, r, p, X) \in \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}^N \times \mathbb{S}^N$  and any sufficiently regular function  $u$ , the nonlinear functional (Hamiltonian)  $H = H(x, r, p, X, u(\cdot))$  is defined by

$$(1.2) \quad H(x, r, p, X, u(\cdot)) = \sup_{\theta \in \Theta} \left\{ -\frac{1}{2} \text{tr}[a^\theta(x)X] - b^\theta(x) \cdot p - \mathcal{J}^\theta u(x) + c^\theta(x)r - f^\theta(x) \right\},$$

where  $a^\theta(x) = \sigma^\theta(x)\sigma^\theta(x)^\top$ , while the integro operator  $\mathcal{J}^\theta$  is defined by

$$(1.3) \quad \mathcal{J}^\theta u(x) = \int_E [u(x + \eta^\theta(x, z)) - u(x) - \mathbf{1}_{|z| < 1} \eta^\theta(x, z) Du(x)] \nu(dz).$$

We use  $\mathbb{S}^N$  to denote the set of symmetric  $N \times N$  matrices  $X = (X_{i,j})$ ,  $i, j = 1, \dots, N$  and  $\Theta$  is a closed subset of Euclidean space. We assume that the  $\mathbb{R}^{N \times P}$ -valued function  $\sigma^\theta$ , the  $\mathbb{R}^N$ -valued functions  $b^\theta$ ,  $\eta^\theta$ , and the  $\mathbb{R}$ -valued functions  $c^\theta(x)$ ,  $f^\theta(x)$  are sufficiently regular (precise conditions will be provided later, but roughly speaking we require that the involved functions are bounded and Lipschitz continuous, uniformly in  $\theta$ ). In (1.3),  $\nu(dz)$  is a given Radon measure on

$$E = \mathbb{R}^M \setminus \{0\},$$

the so-called Lévy measure, which typically possesses a second order singularity at the origin and some exponential decay property at infinity. Due to the second order singularity, it is often useful to split  $\mathcal{J}^\theta u(x)$  as follows:

$$\begin{aligned} \mathcal{J}^\theta u(x) &= \int_0^1 (1 - \rho) \left( \int_{E \cap |z| < \kappa} D^2 v(x + \rho \eta^\theta(x, z)) \eta^\theta(x, z) \eta^\theta(x, z)^\top \nu(dz) \right) d\rho \\ &\quad + \int_{E \cap |z| \geq \kappa} [v(x + \eta^\theta(x, z)) - v(x) - \eta^\theta(x, z) Du(x)] \nu(dz), \end{aligned}$$

which reveals that the integro operator has a component that behaves like a “non-local” second order differential operator with an “ellipticity coefficient” determined by the jump amplitude  $\eta^\theta$  and the Lévy measure  $\nu(dz)$ .

We use the term degenerate elliptic when we speak about (1.1) since the diffusion matrix  $a^\theta(x)$  is merely nonnegative definite. Notice also that the “ellipticity” of  $\mathcal{J}^\theta$  may be degenerate, as the jump amplitude  $\eta^\theta$  in our setup is allowed to be zero. In view of our assumptions (diffusion matrix  $a^\theta$  and jump amplitude  $\eta^\theta$  can be zero), we cannot in general expect the existence of smooth solutions to (1.1). Due to the fully nonlinear nature of the equation (1.1), it is natural to work within the framework of viscosity solutions. Since the theory of viscosity solutions for fully nonlinear degenerate partial differential equations is highly developed [5, 6, 23, 31], there has been an interest in recent years to apply this theory to integro-partial differential equations [1, 2, 3, 4, 7, 14, 15, 16, 38, 39, 49, 50, 51, 53, 54, 56, 57]. Of particular relevance to the present paper are the works [38, 39]. Let us also mention the books [32, 33], which treat linear equations by the method of Green functions.

Nonlocal integro-partial differential equations such as (1.1) arise when one attempts to solve stochastic optimal control problems with the dynamic programming approach [31]. Examples include various types of portfolio optimization problems in which the risky assets follow jump-diffusion (Lévy) processes possessing discontinuous sample paths, see for example [14, 15, 16, 50] and the references cited therein. The value function of such a control problem is the unique viscosity solution of a Bellman equation of the form (1.1). The regularity/growth properties of the value function depend on the properties of the coefficients of the equation. Herein we

work in the class of bounded and Lipschitz continuous viscosity solutions, which is consistent with the conditions discussed above.

When  $\Theta$  is a singleton the equation (1.1) reduces to a linear integro-partial differential equation. Linear integro-partial differential equations are relevant in the pricing theory of financial derivatives. For more information about pricing theory and linear integro-partial differential equations, we refer to [21].

The standard model for the price evolution of stocks (and other risky assets) is the geometric Brownian motion, which assumes that the stock returns are normally distributed. However, it turns out that the normal distribution poorly fits the stock returns, at least on sufficiently small time scales. Indeed, the returns distributions are, for example, leptokurtic and have longer and fatter tails than the normal distribution. To improve upon the shortcomings of the standard model, many jump-diffusion models have been suggested in the literature over the years. Although we will discuss briefly an example later on in this paper, we refer to [21, 55] for information about the most popular models. The empirical fact that Lévy processes with discontinuous sample paths tend to better model, e.g., stock prices is one main reason for the (renewed) interest in stochastic optimal control of jump-diffusion processes and option pricing in markets based on such processes.

In this paper we are interested in approximation schemes for nonlocal equations such as (1.1) and their convergence properties. The approximation schemes that we have in mind are those based on the finite difference method. As a general introduction to finite difference (Markov chain) schemes for stochastic control problems, we refer [43], see also [17, 27, 31, 46]. The main idea behind Markov chain schemes is to replace the original controlled stochastic process by a controlled discrete-time/discrete-space Markov chain. The value function of the discrete control problem satisfies a discrete Bellman equation that can be solved by various fixed point iterations. Let us also mention that there are other classes of approximations schemes for stochastic control problems, see for example [20] for the so-called control schemes.

There are few works on numerical schemes for fully nonlinear degenerate integro-partial differential equations, but see the discussion about jump-diffusion processes in [43]. Furthermore, in [29, 30] a finite difference (Markov chain) scheme along the lines of [43] is proposed and shown to converge to the unique viscosity solution for a singular stochastic optimal control problem of investment-consumption type, with the underlying risky asset being based on an exponential second order singular Lévy model. The construction and analysis of numerical schemes for linear integro-partial differential equations arising as pricing equations in financial markets of jump-diffusion type is currently an active field of research, see the recent papers [18, 19, 21, 22, 26, 45].

Although there is a well-established theory for proving convergence of approximate (viscosity) solutions to second order equations [5, 6, 10, 23, 31], see also [43] for probabilistic proofs, the question of error estimates is much more difficult and remained open until the recent works by Krylov [40, 41, 42] and Barles and Jakobsen [8, 9, 35] on finite difference schemes and more generally on monotone approximation schemes for (convex) Bellman equations associated with controlled diffusions. Herein we use these recent advances for second order partial differential equations to derive error estimates for some approximation schemes for integro-partial differential equations.

In passing, we mention that it is well-known how to derive error estimates in the first order local (convex/nonconvex) Hamilton-Jacobi case, see [24] for the first result in that direction and, e.g., [23] for further references.

Our interest herein is to derive error estimates for monotone approximation schemes for the integro-partial differential equation (1.1). A general approximation scheme can be written abstractly as

$$(1.4) \quad S(h, x, u_h(x), [u_h]_x) = 0, \quad x \in \mathbb{R}^N,$$

where  $h = (h_1, h_2) \in \mathbb{R}^N \times \mathbb{R}^M$  is a collective symbol for the discretization parameters;  $h_1 > 0$  is the  $x$  - mesh size used for the approximation of the partial derivatives in (1.1), while  $h_2 > 0$  is the  $z$  - mesh size used for the approximation of the integro operator in (1.1),  $u_h : \mathbb{R}^N \rightarrow \mathbb{R}$  is the solution of the scheme (1.4), and  $[u_h]_x$  is a function defined at  $x$  in terms of (possibly all) values of  $u_h$ . Roughly speaking, (1.4) is a consistent, monotone ( $S$  is nondecreasing in the  $u_h(x)$  - slot and nonincreasing in the  $[u_h]_x$  - slot), and uniformly continuous approximation of (1.1) (precise conditions will be given in the next section).

Our first main result is an explicit estimate for the pointwise difference between the unique viscosity solution  $u$  of (1.1) and the solution  $u_h$  of the scheme (1.4). To prove the error estimate we use the classical idea of considering a sequence of smooth approximate solutions  $u_\varepsilon$  of (1.1); if such a sequence exists and one has a bound on the derivatives of  $u_\varepsilon$  and  $\|u - v_\varepsilon\|_{L^\infty(\mathbb{R}^N)}$ , then the consistency condition and comparison properties of the scheme allow us to plug  $u_\varepsilon$  into  $S$  to get an estimate for  $\|u_\varepsilon - u_h\|_{L^\infty(\mathbb{R}^N)}$ . These results immediately imply an estimate of  $\|u - u_h\|_{L^\infty(\mathbb{R}^N)}$  depending on  $h$  and  $\varepsilon$ . Optimizing with respect to  $\varepsilon$  we obtain the desired rate of convergence. However, it is not easy to implement this strategy, as it is well explained in [8, 9, 40, 41, 42]. Following an idea Krylov, which he refers to as the method of “shaking the coefficients”, we introduce the auxiliary equation

$$(1.5) \quad \max_{|e| \leq \varepsilon} [H(x + e, u^\varepsilon, Du^\varepsilon, D^2u^\varepsilon, u^\varepsilon(\cdot))] = 0 \text{ in } \mathbb{R}^N.$$

Taking into account the convexity of  $H$ , we then regularize via convolution the unique viscosity solution  $u^\varepsilon$  of (1.5) to obtain a sequence of subsolutions  $u_\varepsilon$  of (1.1). With this and the consistency/comparison principle for the scheme, we can produce an upper estimate for  $u - u_h$ . To obtain the lower estimate, we follow the idea of Barles and Jakobsen [8]: We exchange the role of the problem and of the scheme, and introduce the auxiliary scheme

$$(1.6) \quad \max_{|e| \leq \varepsilon} [S(h, x + e, u_h^\varepsilon(x), [u_h^\varepsilon]_x)] = 0,$$

and then proceed in a similar way as for the upper bound.

Having established an error estimate for a general class of approximation schemes of the form (1.4), an application is presented next, in which we apply it to a finite difference scheme based on upwind (one-sided) differencing of the first order derivatives, central differencing of the second order derivatives, and an integration rule to approximate the integro operator. In this application we assume that  $\sigma^\theta$  does not depend on  $x$  and that the Lévy measure  $\nu(dz)$  is bounded and compactly supported. The existence and uniqueness of a solution to the numerical scheme follow from standard fixed point considerations.

In the general case in which the Lévy measure can be unbounded as well as have unbounded support, we do not directly discretize the integro operator as it stands in (1.1). In particular, due to the second order singularity of the Lévy measure, we follow [29, 30] and replace (1.1) by an approximate equation in which this singularity has been replaced by an additional diffusion term, accounting for the infinite activity of the Lévy process at the origin. To be more precise, for  $r > 0$  small and  $R > 0$  large we introduce a truncated domain  $\{z : r < |z| < R\}$  and a truncated Lévy measure  $\nu_{r,R}(dz)$ :

$$\nu_{r,R}(dz) = \mathbf{1}_{r < |z| < R} \nu(dz),$$

where  $\mathbf{1}$  denotes the characteristic function and the truncated domain not only accounts for the unboundedness at the origin of the Lévy measure, but also the unboundedness of the domain  $E$ . The approximate equation then reads

$$(1.7) \quad H_{r,R}(x, u(x), Du(x), D^2u(x), u(\cdot)) = 0 \quad \text{in } \mathbb{R}^N,$$

where the approximate Hamiltonian  $H_{r,R}$  is obtained by replacing  $a^\theta$  and  $\mathcal{J}^\theta$  in (1.2) by  $a^\theta + a_{r,R}^\theta$  and  $\mathcal{J}_{r,R}^\theta$ , respectively. The additional diffusion coefficient  $a_{r,R}^\theta$  is supposed to be chosen such that it is a model for the infinite activity of the original Lévy process at origin (we suggest several choices of  $a_{r,R}^\theta$ ), while  $\mathcal{J}_{r,R}^\theta$  is the integro operator obtained by replacing the Lévy measure  $\nu(dz)$  in (1.3) by the truncated measure  $\nu_{r,R}(dz)$ . As the measure  $\nu_{r,R}(dz)$  is bounded and compactly supported, the integro-partial differential equation (1.7) can be discretized by the finite difference scheme introduced earlier.

To control the error induced by truncating the Lévy measure and by introducing the additional diffusion coefficient, we need to have an estimate on the difference between the viscosity solution  $u$  of the original equation (1.1) and the viscosity solution  $u_{r,R}$  of the perturbed equation (1.7). Such an estimate follows from the recent continuous dependence result proved in [38]. However, it turns out that in some cases we can have better results by using the same approximation method we used to prove the error bounds for the numerical schemes, and this is the approach we take in this paper. A key feature of our results is that they reveal how the choice of  $a_{r,R}^\theta$  influences the accuracy of our approximations.

Finally, let us mention that by combining the methods in this paper with those in [35], we can treat time dependent problems. Moreover, we can include certain  $x$ -dependence in  $\sigma^\theta$  by using the recent methods of Krylov [42], or a completely general  $x$  - dependence in  $\sigma^\theta$  (at the expense of lower convergence rates) following the arguments of [9]. The details will be presented in future work.

The remaining part of this paper is organized as follows: To explain the origin of equations like (1.1), we use Section 2 to say a few words about stochastic optimal control of jump-diffusion processes and give an example of a jump-diffusion process that is particularly popular in finance applications. In Section 3 we introduce some notations to be used throughout this paper, list various conditions regarding the coefficients of the equation (1.1), and recall the definition of viscosity solutions as well as some well-posedness/regularity results for integro-partial differential equations. In Section 4 we provide precise conditions on the approximation scheme (1.4) along with the statement and proof of the main error estimate. In Section 5 we present a finite difference scheme for (1.1) and we verify the conditions needed to apply the error estimate from Section 4. Finally, in Section 7 we prove some technical results used elsewhere in the text.

## 2. OPTIMAL CONTROL OF JUMP-DIFFUSION PROCESSES

In this section we provide some background information on stochastic optimal control, dynamic programming, and Lévy processes. The purpose of this section is to complement Section 1 and briefly explain the origin of nonlinear integro-partial differential equations like (1.1), thereby setting the scene for the whole paper.

**2.1. The optimal control problem.** Let  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$  be a filtered complete probability space satisfying the usual hypotheses. Let  $\Theta, b^\theta, \sigma^\theta, \eta^\theta, f^\theta, c^\theta$  be as discussed in Sections 1 and 3. Let  $\{W_t\}_{t \geq 0}$  be a  $P$  - dimensional Wiener process and  $\{N_t\}_{t \geq 0}$  be a homogeneous Poisson random measure with intensity measure  $\nu(dz) \times dt$ , where  $\nu(dz)$  is the Lévy measure on  $E$ . These processes are independent and adapted to the filtration  $\{\mathcal{F}_t\}_{t \geq 0}$ . The compensated Poisson random measure

of  $N$  is given by  $\tilde{N}(dz, dt) = N(dz, dt) - \nu(dz) \times dt$ . Let  $\{\theta_t\}_{t \geq 0}$  be an adapted control process taking values in  $\Theta$ . Consider then the (controlled) stochastic differential equation

$$(2.1) \quad dX_t = b^{\theta_t}(X_t) dt + \sigma^{\theta_t}(X_t) dW_t + \int_E \eta^{\theta_t}(X_{t-}, z) \tilde{N}(dz, dt), \quad t > 0.$$

Under the conditions on  $b^\theta, \sigma^\theta, \eta^\theta$  stated in Section 3, there exists a unique solution

$$\left\{ X_t = X_t^{\theta, x} \right\}_{t \geq 0}$$

of (2.1) with initial condition  $X_t = x$ .

Given  $x \in \mathbb{R}^N$  and an adapted control process  $\theta = \{\theta_t\}_{t \geq 0}$ , the infinite horizon stochastic optimal control problem is to minimize the functional

$$(2.2) \quad J(x; \theta) = \mathbb{E}^{x, \theta} \left[ \int_0^\infty f^{\theta_s}(X_s) \exp \left( - \int_0^s c^{\theta_r}(X_r) dr \right) ds \right].$$

To solve this optimization problem with the dynamic programming method (see [31]) we introduce the value function

$$(2.3) \quad V(x) = \inf_{\theta \in \Theta} J(x; \theta), \quad x \in \mathbb{R}^N.$$

Under the conditions on  $b^\theta, \sigma^\theta, \eta^\theta, f^\theta, c^\theta$  listed in Section 3, the value function  $V(x)$  is bounded and Lipschitz continuous on  $\mathbb{R}^N$ , see [31, 48, 47, 51].

At the heart of the matter is the dynamic programming principle (see [31]): for any stopping time  $\tau$  and  $t \geq 0$ ,

$$V(x) = \inf_{\theta \in \Theta} \mathbb{E}^{x, \theta} \left[ \int_0^{t \wedge \tau} f^{\theta_s}(X_s) \exp \left( - \int_0^s c^{\theta_r}(X_r) dr \right) ds + V(X_{t \wedge \tau}) \exp \left( - \int_0^{t \wedge \tau} c^{\theta_r}(X_r) dr \right) \right],$$

where  $t \wedge \tau = \min(t, \tau)$ . A well-known consequence of this principle, which roughly speaking says that “global optimality” implies “local optimality”, is that the value function (2.3) is a viscosity solution of the Bellman equation (1.1), see for example Pham [51].

**2.2. Lévy processes.** We recall that a stochastic process  $L = \{L_t\}_{t \geq 0}$  is called a Lévy process if  $L_0 = 0$ ,  $L$  has independent and stationary increments, and  $L$  is continuous in probability. For simplicity dropping the control dependence and assuming that  $N = M = 1$ , so that in particular  $x \in \mathbb{R}$  and  $z \in E = \mathbb{R} \setminus \{0\}$ , we may consider a Lévy processes of the form

$$(2.4) \quad dL_t = \mu dt + \sigma dW_t + \int_{|z| < 1} \eta z \tilde{N}(dz, dt) + \int_{|z| \geq 1} \eta z N(dz, dt),$$

where  $\mu, \eta \in \mathbb{R}, \sigma > 0$ , and the Wiener process  $W = \{W_t\}_{t \geq 0}$ , the Poisson random measure  $N = \{N_t\}_{t \geq 0}$ , and the compensated Poisson random measure  $\tilde{N} = \{\tilde{N}_t\}_{t \geq 0}$  have been introduced above. According to the Lévy-Khintchine representation, any Lévy process can be written in a form similar to (2.4) [21, 55].

Let us consider a financial market where the risk-free asset  $B_t$  and the risky asset  $X = \{X_t\}_{t \geq 0}$  evolve according to the exponential models

$$(2.5) \quad B_t = e^{rt}, \quad X_t = e^{L_t},$$

where  $r > 0$  is the interest rate. By the generalized Itô's formula [21, 55], the  $X_t$  dynamics are given by (2.1) (without the control dependence and with  $x \in \mathbb{R}, z \in E = \mathbb{R} \setminus \{0\}$ ) with

$$\begin{aligned} b(X_t) &= \left[ \mu + \frac{1}{2}\sigma^2 + \int_E \left( e^{\eta z} - 1 - \eta z \mathbf{1}_{|z| < 1} \right) \nu(dz) \right] X_t, \\ \sigma(X_t) &= \sigma X_t, \quad \eta(X_{t-}, z) = [e^{\eta z} - 1] X_{t-}. \end{aligned}$$

Exponential Lévy models like the one in (2.5) have been used recently in some portfolio optimization problems in [13, 14, 15, 16], see [4, 21] and the references therein for the use of such models in option pricing theory.

Market models like (2.5) are more realistic alternatives to the standard Black-Scholes model in which  $X_t = x e^{\mu t + \sigma W_t}$ , where  $\mu$  is the drift coefficient (mean rate of return),  $\sigma$  is the volatility, and  $W = \{W_t\}_{t \geq 0}$  is a Brownian motion. This model produces logreturns which are normally distributed with mean  $\mu \Delta t$  and variance  $\sigma^2 \Delta t$ , where  $\Delta t$  denotes the chosen time window, measured for example in weeks, days, minutes, etc. However, as was discussed in the introduction, empirical studies of logreturn data from for example stock markets show large deviation from normality when  $\Delta t$  is small (e.g., daily), which is a fact that has motivated many authors to study various problems related to optimal control, option pricing, etc., in which the market is modeled by an exponential Lévy process (2.5).

As already mentioned in the introduction, many different Lévy models (i.e., choices of  $L$  in (2.5)) have been suggested in the literature over the years. We refer to [21, 55] for an overview of this activity. Here we mention the HYP (hyperbolic) Lévy model proposed in [28] as a model for German stock prices, which was shown to give a good fit. Another example is the NIG (normal inverse Gaussian) Lévy model suggested in [12], which was shown to model German stock prices well in [52]. The two mentioned models belong to the class of GH (generalized hyperbolic) Lévy models. They are characterized by having independent increments belonging to the class of GH distributions. This class of distributions, and in particular its two corresponding subclasses of NIG distributions and HYP distributions, has proved to provide a good fit to empirically observed log-returns. The class of GH distributions [11] can be characterized as normal variance-mean mixtures, with the mixing distribution being a GIG distribution. This class of distributions includes many interesting subclasses, and limiting cases like the NIG, HYP, VG, Student- $t$ , and normal distributions.

The density function of a GH distribution, let us call it  $f^{\text{GH}}$ , depends on five parameters

$$\lambda \in \mathbb{R}, \quad \alpha > 0, \quad \beta \in (-\alpha, \alpha), \quad \delta > 0, \quad \mu \in \mathbb{R},$$

with the following interpretation:  $\alpha$  is a steepness parameter (larger  $\alpha$  means steeper density),  $\beta$  is a parameter of asymmetry (if  $\beta = 0$  the density is symmetric around the mean),  $\delta$  is a scale parameter, and  $\mu$  is a location parameter. The special case of  $\lambda = -\frac{1}{2}$  gives a NIG distribution. For  $\lambda = \frac{1}{2}$  we get a HYP distribution. The Lévy measure  $\nu(dz) = \nu^{\text{GH}}(dz)$  of a GH distribution is absolutely continuous with respect to the Lebesgue measure  $dz$ . As an example, let us display the density function of the NIG distribution (a subclass of the GH distributions):

$$f^{\text{NIG}}(y) = \frac{\alpha \delta}{\pi} \exp\left(\delta \sqrt{\alpha^2 - \beta^2} + \beta(y - \mu)\right) \frac{K_1(\alpha \sqrt{\delta^2 + (y - \mu)^2})}{\sqrt{\delta^2 + (y - \mu)^2}}, \quad y \in \mathbb{R}$$

where  $K_1$  is the modified Bessel function of the third kind and index 1, i.e.,

$$K_1(y) = \frac{1}{2} \int_0^\infty \exp\left(-\frac{1}{2}y(s + s^{-1})\right) ds, \quad \text{for } y > 0,$$

where  $z \in \mathbb{R}$ ,  $\mu \in \mathbb{R}$ ,  $\delta > 0$ , and  $0 \leq |\beta| \leq \alpha$ . In this case the Lévy measure takes the form

$$\nu^{\text{NIG}}(dz) = \nu^{\text{NIG}}(z) dz = \frac{\alpha\delta}{\pi|z|} e^{\beta z} K_1(\alpha|z|) dz.$$

Due to the properties of the Bessel function, the Lévy density  $\nu^{\text{NIG}}(z)$  behaves like  $1/|z|^2$  near the origin. The Lévy-Khintchine formula for the normal inverse Gaussian Lévy process takes the form

$$L_t = \xi t + \int_0^t \int_{\mathbb{R} \setminus \{0\}} z \tilde{N}(dt, dz), \quad \xi = \mu + \frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}}.$$

### 3. PRELIMINARIES

In this section we will explain our notation, list the main assumptions, define what we mean by a (viscosity) solution of (1.1), and state some well-posedness results for such solutions.

For  $\mathcal{O} \subset \mathbb{R}^N$  we denote by  $USC(\mathcal{O})$ ,  $LSC(\mathcal{O})$ ,  $C(\mathcal{O})$ ,  $C^{0,1}(\mathcal{O})$ ,  $C^n(\mathcal{O})$ ,  $C^\infty(\mathcal{O})$  respectively the spaces of upper semicontinuous functions, lower semicontinuous functions, continuous functions, bounded Lipschitz continuous functions,  $n$  times continuously differentiable functions, and infinitely differentiable functions on  $\mathcal{O}$ . The subscript “ $b$ ” will denote bounded functions and the subscript “ $\text{pol}$ ” will denote polynomial growth at infinity. For example if  $f : \mathcal{O} \rightarrow \mathbb{R}$  belongs to  $C_{\text{pol}}(\mathcal{O})$ , then there is a  $m \in \mathbb{N}$  such that

$$|f(x)| \leq C(1 + |x|^m), \text{ for all } x \in \mathcal{O}.$$

We will use the following norms

$$|f|_0 = \sup_{x \in \mathbb{R}^N} |f(x)|, \quad [f]_1 = |Df|_0, \quad \text{and} \quad |f|_1 = |f|_0 + [f]_1.$$

In general we will denote by  $D^i f$ , for  $i \in \mathbb{N}$ , the vector of the  $i$ th order partial derivatives of  $f$ . For  $r > 0$  and  $x \in \mathbb{R}^N$  we define  $B(x, r) = \{y \in \mathcal{O} : |y - x| < r\}$ . Moreover, for  $0 < \kappa < 1$  we define

$$(3.1) \quad E_\kappa = \left\{ z \in \mathbb{R}^M : 0 < |z| < \kappa \right\}, \quad \text{and} \quad E^\kappa = \mathbb{R}^M - E_\kappa.$$

We let  $\mathbf{1}_{\mathcal{O}}(z)$  denote the characteristic function of the set  $\mathcal{O}$ , and finally, for  $\varepsilon > 0$  we define a sequence of mollifiers  $(\rho_\varepsilon)_\varepsilon$  as

$$(3.2) \quad \rho_\varepsilon(x) = \frac{1}{\varepsilon^N} \rho\left(\frac{x}{\varepsilon}\right),$$

where  $\rho$  is a positive  $C^\infty(\mathbb{R}^N)$  function with unit mass and support in  $B(0, 1)$ .

In this paper we will use the following assumptions (cf. equation (1.1)):

(A.1)  $\nu$  is a positive radon measure on  $E$  such that

$$\int_{E_1} |z|^2 \nu(dz) + \int_{E^1} e^{\ell|z|} \nu(dz) < \infty$$

for some given  $\ell > 0$ , where  $E_1$  and  $E^1$  are defined above.

(A.2) There exists  $M > 0$  such that for any  $\theta \in \Theta$

$$\begin{aligned} |\sigma^\theta|_1 + |b^\theta|_1 + |c^\theta|_1 + |f^\theta|_1 &\leq M, \\ |\eta^\theta(\cdot, z)|_1 &\leq M \left( |z| \mathbf{1}_{E_1}(z) + e^{\ell|z|} \mathbf{1}_{E^1}(z) \right), \end{aligned}$$

where  $\mathbf{1}_{E_1}, \mathbf{1}_{E^1}$  are characteristic functions defined above.



(A.3') There exists  $\lambda > 0$  such that, for all  $x \in \mathbb{R}$  and for all  $\theta \in \Theta$ ,

$$c^\theta(x) \geq \lambda.$$

Under assumption (A.1) the measure  $\nu$  may have a singularity at  $z = 0$ . Furthermore, under (A.1), the integro-term in (1.1), and hence also the equation itself, is well-defined in the class of  $C_{\text{pol}}^2$  functions. This follows from the exponential decay at infinity of  $\nu$  and the fact that the singularity at  $z = 0$  is integrable for any  $C^2$  function  $u$  (to see this do a second order Taylor expansion). It is well known that under the above assumptions the solutions to (1.1) need not be smooth, so that a concept of weak solutions needs to be employed. It turns out that a good concept of weak solutions is that of viscosity solutions.

**Definition 3.1.** A function  $u \in USC(\mathbb{R}^N)$  ( $u \in LSC(\mathbb{R}^N)$ ) is a viscosity subsolution (supersolution) of (1.1) if for every  $x \in \mathbb{R}^N$  and  $\phi \in C_{\text{pol}}^2(\mathbb{R}^N)$  such that  $x$  is a global maximum (minimum) for  $u - \phi$ ,

$$H(x, u(x), D\phi(x), D^2\phi(x), \phi(\cdot)) \leq 0 \quad (\geq 0).$$

Next we want to give an equivalent definition of viscosity solutions. To do that, we split the integral term  $\mathcal{J}^\theta u$  into two parts:

$$(3.3) \quad \mathcal{J}^\theta(u, \phi)(x) = \mathcal{J}_\kappa^\theta \phi(x) + \mathcal{J}^{\theta, \kappa}(u, D\phi(x))(x),$$

where  $0 < \kappa < 1$ ,  $E_\kappa$  and  $E^\kappa$  are defined in (3.1), and

$$\begin{aligned} \mathcal{J}_\kappa^\theta \phi(x) &= \int_{E_\kappa} [\phi(x + \eta^\theta(x, z)) - \phi(x) - \eta^\theta(x, z)D\phi(x)] \nu(dz), \\ \mathcal{J}^{\theta, \kappa}(u, p)(x) &= \int_{E^\kappa} [u(x + \eta^\theta(x, z)) - u(x) - \mathbf{1}_{|z| < 1} \eta^\theta(x, z)p] \nu(dz). \end{aligned}$$

It is easy to see that  $\mathcal{J}_\kappa^\theta \phi$  is well-defined when  $\phi \in C^2(\mathbb{R}^N)$  and that  $\mathcal{J}^{\theta, \kappa}(u, p)$  is well-defined for  $p \in \mathbb{R}^N$  when  $u \in LSC_{\text{pol}}(\mathbb{R}^N) \cup USC_{\text{pol}}(\mathbb{R}^N)$ . In addition, it is easy to establish the following result (see for example [51]):

**Lemma 3.1.** If (A.1) and (A.2) hold, and  $\phi \in C^2(\mathbb{R}^N)$ , then

$$\lim_{\kappa \rightarrow 0} \mathcal{J}_\kappa^\theta(\phi)(x) = 0$$

uniformly in  $x \in \mathbb{R}^N$  and  $\theta \in \Theta$ .

We can rewrite (1.1) as

$$(3.4) \quad H_\kappa(x, u(x), Du(x), D^2u(x), u(\cdot), u(\cdot)) = 0.$$

where

$$\begin{aligned} &H_\kappa(x, \phi, p, X, \phi(\cdot), v(\cdot)) \\ &= \sup_{\theta \in \Theta} \left\{ -\frac{1}{2} \text{tr}[\sigma^\theta \sigma^{\theta^\top} X] - b^\theta(x)p - \mathcal{J}_\kappa^\theta \phi(x) - \mathcal{J}^{\theta, \kappa}(v, p)(x) + c^\theta(x)\phi - f^\theta(x) \right\}. \end{aligned}$$

Now we state an equivalent definition of viscosity solutions (see for example [51]).

**Lemma 3.2** (Alternative definition). Assume (A.1), (A.2), and (A.3') hold. A function  $u \in USC_{\text{pol}}(\Omega)$  ( $u \in LSC_{\text{pol}}(\Omega)$ ) is a viscosity subsolution (supersolution) of (1.1) if and only if for every  $x \in \mathbb{R}^N$  and  $\phi \in C^2(\mathbb{R}^N)$  such that  $x$  is a global maximizer (minimizer) for  $u - \phi$ ,

$$H_\kappa(x, u(x), D\phi(x), D^2\phi(x), u(\cdot), \phi(\cdot)) \leq 0 \quad (\geq 0) \text{ for every } \kappa \in (0, 1).$$

With the use of this equivalent definition and the maximum principle for semi-continuous functions [23], adapted to integro-PDEs in [39], it is standard to prove existence, uniqueness, and regularity results, see, e.g., [4, 7, 39, 38, 51]. Here we state such results without proofs.

**Theorem 3.3.** *Assume (A.1), (A.2), and (A.3') hold.*

(i) *There exists a unique viscosity solution  $u \in C_b(\mathbb{R}^N)$  of equation (1.1) which is Hölder continuous, i.e., there is a  $\delta \in (0, 1]$  such that*

$$|u(x) - u(y)| \leq C|x - y|^\delta \quad \text{for all } x, y \in \mathbb{R}^N.$$

(ii) *There exists a constant  $\lambda_0$  depending only on  $N$ ,  $\sup_\theta[\sigma^\theta]_1$ ,  $\sup_\theta[b^\theta]_1$ , and  $\sup_\theta \int_E [\eta^\theta(\cdot, z)]_1^2 \nu(dz)$  such that if*

$$\lambda \geq \lambda_0,$$

*then the viscosity solution  $u$  of (1.1) is Lipschitz continuous ( $\delta = 1$  above).*

(iii) *Let  $u, -v \in USC_b(\mathbb{R}^N)$ . If  $u$  and  $v$  are respectively viscosity sub- and supersolutions of (1.1), then*

$$u \leq v \quad \text{in } \mathbb{R}^N.$$

In the remaining part of this paper we always make the following assumption:

**(A.3)** For all  $x \in \mathbb{R}$  and for all  $\theta \in \Theta$ ,

$$c^\theta(x) \geq \lambda_0,$$

where  $\lambda_0$  is defined in Theorem 3.3 (ii).

Hence, in this paper the viscosity solutions to (1.1) are always Lipschitz continuous. However, the results in this paper extend easily to the general case (any  $\lambda > 0$  and Hölder continuous solutions) by approximation - we refer to [8] where this is explained in the pure diffusion case.

#### 4. AN ABSTRACT FRAMEWORK FOR OBTAINING ERROR BOUNDS

In this section we develop an abstract framework for deriving error bounds for approximate solutions to integro-partial differential equations like (1.1). We begin by stating and explaining a set of conditions that have to be imposed on the general approximation scheme (1.4). Then we state and prove the main error bound.

We impose the following conditions on our approximation scheme (1.4):

**(C.1)** (*Monotonicity of the scheme*) There exists  $\bar{\lambda} > 0$  such that, for every  $h \geq 0$ ,  $x \in \mathbb{R}^N$ ,  $\zeta \in \mathbb{R}$ ,  $\mu \geq 0$ , and bounded function  $u, v$  such that  $u \leq v$  in  $\mathbb{R}^N$ , then

$$S(h, x, \zeta + \mu, [u + \mu]_x) \geq S(h, x, \zeta, [v]_x) + \bar{\lambda}\mu.$$

**(C.2)** (*Regularity*) For every  $h > 0$  and  $\phi \in C_b(\mathbb{R}^N)$ , the function

$$x \mapsto S(h, x, \phi(x), [\phi]_x)$$

is bounded and continuous on  $\mathbb{R}^N$  and the function

$$\zeta \mapsto S(h, x, \zeta, [\phi]_x),$$

is uniformly continuous for bounded  $\zeta$ , uniformly with respect to  $x \in \mathbb{R}^N$ .

**(C.3)** (*Consistency*) There exist  $n \in \mathbb{N}$ ,  $K_i, \bar{K}_i, k_i, \bar{k}_i \geq 0$ ,  $i = 1, \dots, n$ , such that for every  $v \in C^n(\mathbb{R}^N)$ ,  $h \geq 0$ , and  $x \in \mathbb{R}^N$

$$|H(x, v, Dv, D^2v, v(\cdot)) - S(h, x, v(x), [v]_x)| \leq \sum_{i=1}^n |D^i v|_0 \left( K_i h_1^{k_i} + \bar{K}_i h_2^{\bar{k}_i} \right).$$

(C.4) (*Convexity*) Let  $(\rho_\varepsilon)_{\varepsilon>0}$  be the mollifiers defined in Section 3. For every  $v \in C^{0,1}(\mathbb{R}^N)$ , there exists a constant  $K > 0$  such that for every  $x$  and  $h$

$$\int_{\mathbb{R}^N} S(h, x, v(x-e), [v(\cdot-e)]_x) \rho_\varepsilon(e) de \geq S(h, x, (v * \rho_\varepsilon)(x), [v * \rho]_x) - K\varepsilon.$$

(C.5) (*Commutation with translations*) For any  $h > 0$  small enough,  $0 \leq \varepsilon \leq 1$ ,  $y \in \mathbb{R}^N$ ,  $\zeta \in \mathbb{R}$ ,  $v \in C_b(\mathbb{R}^N)$  and  $|e| \leq \varepsilon$ , we have

$$S(h, y, \zeta, [v]_{y-e}) = S(h, y, \zeta, [v(\cdot-e)]_y).$$

Assumption (C.4) is natural here because (1.1) is convex. By Jensen's inequality, it holds for any scheme (1.4) which is convex in  $(u_h(x), [u_h]_x)$  [8].

**Definition 4.1.** A function  $u \in C_b(\mathbb{R}^N)$  is a subsolution (supersolution) of the scheme (1.4) if

$$S(h, x, u(x), [u]_x) \leq 0 (\geq 0) \quad \text{in } \mathbb{R}^N.$$

From assumptions (C.1) and (C.2) we can derive a comparison result for the scheme (1.4) (see [44, 8]) implying uniqueness in the  $C_b(\mathbb{R}^N)$  class of solutions.

**Lemma 4.1.** Assume (C.1) and (C.2) hold and let  $u, v \in C_b(\mathbb{R}^N)$ . If  $u$  and  $v$  are sub- and supersolutions of (1.4) respectively, then  $u \leq v$  in  $\mathbb{R}^N$ .

Before we state the final assumption on the scheme, we will try to explain why it appears. To this end, we outline first the proof of an upper bound on the error. We use the regularization procedure discussed in the introduction. Following a clever idea by Krylov [41], we introduce the intermediate equation (1.5), which has the same structure as equation (1.1) (consider  $(\theta, e)$  as the new control parameters), and the following existence, uniqueness, and convergence results hold:

**Lemma 4.2.** Assume (A.1) - (A.3) hold and let  $0 \leq \varepsilon \leq 1$ . Then there exists a unique viscosity solution  $u^\varepsilon \in C^{0,1}(\mathbb{R}^N)$  to (1.5) satisfying

$$|u^\varepsilon|_1 \leq C \quad \text{and} \quad |u^\varepsilon - u|_0 \leq C\varepsilon,$$

where the constant  $C$  is independent of  $\varepsilon$ .

The existence, uniqueness, and regularity part of this lemma follows directly from Theorem 3.3, while the error estimate follows from Theorem 7.2 in Section 7.

We mollify  $u^\varepsilon$  and prove that  $u_\varepsilon := u^\varepsilon * \rho_\varepsilon$  is a viscosity subsolution of (1.1):

**Lemma 4.3.** Assume (A.1) - (A.3) hold. Then the function  $u_\varepsilon$  defined above is a viscosity subsolution of (1.1).

This result follows from Theorem 7.4 (cf. Section 7), see also the proof below for similar arguments. Using the properties of  $u_\varepsilon$  and a classical argument using consistency and monotonicity of the scheme, followed by a minimization over  $\varepsilon$ , then leads to the upper bound on  $u - u_h$  (see the proof of Theorem 4.4 below).

To conclude we also need a lower bound on  $u - u_h$ . To get this bound, the idea is to reverse the roles of  $u$  and  $u_h$  [8, 40]. So we introduce the intermediate scheme (1.6). Now it turns out that assumptions (A.1) - (A.3) and (C.1) - (C.5) are not strong enough to ensure existence, uniqueness, and convergence results for  $u_h^\varepsilon$  similar to Lemma 4.2. Instead we will simply impose the existence of such results as a condition on the approximation scheme (1.6):

- (D) For  $h > 0$  small enough and  $0 \leq \varepsilon \leq 1$ , the scheme (1.6) has a solution  $u_h^\varepsilon \in C^{0,1}(\mathbb{R}^N)$ . Moreover there is a constant  $C$ , independent of  $h$  and  $\varepsilon$ , such that

$$|u_h^\varepsilon|_1 \leq C \text{ and } |u_h^0 - u_h^\varepsilon| \leq C\varepsilon,$$

where  $u_h^0$  is the solution of (1.4).

This assumption has to be checked every time a particular scheme is studied. We will do so for a class of finite difference schemes in Section 5. If the second order coefficient  $\sigma$  in (1.1) is independent of  $x$ , this assumption probably holds for most schemes of interest satisfying (C.1) - (C.5), and this is the case we will consider in Section 5. The case when  $\sigma$  depends on  $x$  is much harder and will be considered in a future work. For pure diffusions there are results for so-called control schemes [8] and for a class of finite difference schemes [42]. We also mention [9] (see also [41]) where assumption (D) is avoided altogether, and results are obtained for any consistent, stable, and monotone scheme (again in the pure diffusion case).

Now we state the main error estimate for the abstract scheme (1.4).

**Theorem 4.4.** *Assume (A.1) - (A.3) and (C.1) - (C.3) hold. Let  $u$  be the viscosity solution of (1.1) and  $u_h$  the solution of the scheme (1.4).*

(a) *Then the following upper bound on the error holds:*

$$u - u_h \leq C(h_1^\delta + h_2^{\bar{\delta}}) \quad \text{in } \mathbb{R}^N,$$

where

$$\delta = \min_{i:K_i>0} \frac{k_i}{i} \quad \text{and} \quad \bar{\delta} = \min_{i:\bar{K}_i>0} \frac{\bar{k}_i}{i},$$

and  $K_i, \bar{K}_i, k_i, \bar{k}_i$  are defined in (C.3).

(b) *If in addition (C.4), (C.5), and (D) hold, then the following lower bound on the error holds:*

$$-C(h_1^\delta + h_2^{\bar{\delta}}) \leq u - u_h \quad \text{in } \mathbb{R}^N.$$

*Remark 4.5.* For (D) to hold, we need that  $\bar{\lambda} > \bar{\lambda}_0$  for some  $\bar{\lambda}_0 > 0$ . This is similar to the situation for equation (1.1), see assumption (A.3) in Section 3. Warning: If either  $\lambda < \lambda_0$  or  $\bar{\lambda} < \bar{\lambda}_0$ , and

$$\frac{\lambda}{\lambda_0} \neq \frac{\bar{\lambda}}{\bar{\lambda}_0},$$

then we will get *different* exponents in the upper and lower bounds on the error! We refer to [8] where this point is explained in the pure diffusion case.

The proof is an extension of the proof given in [8] (see also [40]).

*Proof.* We start by proving the the upper bound on  $u - u_h$ .

As we have seen above, by Lemmas 4.2 and 4.3  $u_\varepsilon = \rho_\varepsilon * u^\varepsilon$  is a subsolution of (1.1). By Lemma 4.2 and properties of mollifiers it follows that  $u_\varepsilon \in C^\infty(\mathbb{R}^N)$ ,  $|u_\varepsilon|_{0,1} \leq C$ , and  $|D^i u_\varepsilon|_0 \leq C\varepsilon^{1-i}$ . By (C.3) we then have

$$\begin{aligned} & H(x, u_\varepsilon, Du_\varepsilon, D^2 u_\varepsilon, u_\varepsilon(\cdot)) - S(h, x, u_\varepsilon(x), [u_\varepsilon]_x) \\ & \geq - \sum_{i=1}^n |D^i u_\varepsilon|_0 \left( K_i h_1^{k_i} + \bar{K}_i h_2^{\bar{k}_i} \right) \\ & \geq -C \sum_{i=1}^n \varepsilon^{1-i} \left( K_i h_1^{k_i} + \bar{K}_i h_2^{\bar{k}_i} \right) =: C(h_1, h_2, \varepsilon), \end{aligned}$$

for all  $x \in \mathbb{R}^N$ , and Lemma 4.3 now yields

$$S(h, x, u_\varepsilon(x), [u_\varepsilon]_x) \leq C(h_1, h_2, \varepsilon) \quad \text{in } \mathbb{R}^N.$$

From (C.1) it follows that  $u_\varepsilon - C(h_1, h_2, \varepsilon)/\bar{\lambda}$  is a subsolution of (1.6), and therefore by the comparison principle for the scheme (Lemma 4.1) we get

$$u_\varepsilon - u_h \leq C(h_1, h_2, \varepsilon)/\bar{\lambda} \quad \text{in } \mathbb{R}^N.$$

From Lemma 4.2 and properties of mollifiers we have

$$|u - u_\varepsilon|_0 \leq C\varepsilon.$$

We conclude that

$$u - u_h \leq C(h_1, h_2, \varepsilon)/\bar{\lambda} + C\varepsilon \quad \text{in } \mathbb{R}^N,$$

and the upper bound follows after a minimization over  $\varepsilon$ .

The proof of the lower bound on  $u - u_h$  is derived similarly: the key idea is to interchange the role of the scheme and the equation.

Let  $u_h^\varepsilon \in C^{0,1}(\mathbb{R}^N)$  be the solution of the scheme (1.6) given by assumption (D). Apply the change of variable  $y = x + e$ , and observe that by (C.5) we have

$$(4.1) \quad S(h, y, u_h^\varepsilon(y - e), [u_h^\varepsilon(\cdot - e)]_y) \leq 0 \quad \text{in } \mathbb{R}^N \quad \text{for all } |e| \leq \varepsilon.$$

As before, we consider

$$u_h^\varepsilon * \rho_\varepsilon(x) = \int_{\mathbb{R}^N} u_h^\varepsilon(x - e) \rho_\varepsilon(e) de.$$

Multiply (4.1) by  $\rho_\varepsilon(e)$ , integrate with respect to  $e$ , and use (C.4) to see that

$$(4.2) \quad \begin{aligned} 0 &\geq \int_{\mathbb{R}^N} S(h, y, u_h^\varepsilon(y - e), [u_h^\varepsilon(\cdot - e)]_x) \rho_\varepsilon(e) de \\ &\geq S(h, y, (u_h^\varepsilon * \rho_\varepsilon)(y), [u_h^\varepsilon * \rho_\varepsilon]_y) - K\varepsilon. \end{aligned}$$

From assumption (D) and properties of mollifiers  $u_h^\varepsilon * \rho_\varepsilon \in C^\infty(\mathbb{R}^N)$ ,  $|u_h^\varepsilon * \rho_\varepsilon|_{0,1} \leq C$ , and  $|D^i u_h^\varepsilon * \rho_\varepsilon|_0 \leq C\varepsilon^{1-i}$ . Then (C.3) yields

$$\begin{aligned} &S(h, x, (u_h^\varepsilon * \rho_\varepsilon)(x), [u_h^\varepsilon * \rho_\varepsilon]_x) \\ &\quad - H(x, u_h^\varepsilon * \rho_\varepsilon, D(u_h^\varepsilon * \rho_\varepsilon), D^2(u_h^\varepsilon * \rho_\varepsilon), (u_h^\varepsilon * \rho_\varepsilon)(\cdot)) \\ &\geq - \sum_{i=1}^n |D^i u_h^\varepsilon * \rho_\varepsilon|_0 \left( K_i h_1^{k_i} + \bar{K}_i h_2^{\bar{k}_i} \right) \\ &\geq -C \sum_{i=1}^n \varepsilon^{1-i} \left( K_i h_1^{k_i} + \bar{K}_i h_2^{\bar{k}_i} \right) =: \bar{C}(h_1, h_2, \varepsilon), \end{aligned}$$

so by (4.2) we see that

$$H(x, u_h^\varepsilon * \rho_\varepsilon, D(u_h^\varepsilon * \rho_\varepsilon), D^2(u_h^\varepsilon * \rho_\varepsilon), (u_h^\varepsilon * \rho_\varepsilon)(\cdot)) \leq C\varepsilon + \bar{C}(h_1, h_2, \varepsilon) \quad \text{in } \mathbb{R}^N.$$

Assumption (A.3) now implies that  $u_h^\varepsilon * \rho_\varepsilon - (C\varepsilon + \bar{C}(h_1, h_2, \varepsilon))/\lambda$  is a subsolution of (1.1), and the comparison principle (Theorem 3.3 (iii)) yields

$$u_h^\varepsilon * \rho_\varepsilon - u \leq (C\varepsilon + \bar{C}(h_1, h_2, \varepsilon))/\lambda \quad \text{in } \mathbb{R}^N.$$

From (D) and properties of mollifiers we see that

$$|u_h - u_h^\varepsilon * \rho_\varepsilon|_0 \leq C\varepsilon \quad \text{in } \mathbb{R}^N,$$

so we have

$$u_h - u = u_h - u_h^\varepsilon * \rho_\varepsilon + u_h^\varepsilon * \rho_\varepsilon - u \leq C(\varepsilon + \bar{C}(h_1, h_2, \varepsilon)) \quad \text{in } \mathbb{R}^N.$$

Minimizing with respect to  $\varepsilon$  then leads to the desired result.  $\square$

In the next section we apply Theorem 4.4 to a particular finite difference scheme for (1.1), at least under the assumptions that  $\sigma$  is independent of  $x$  and the Lévy measure  $\nu(dz)$  is finite and compactly supported. How to deal with general Lévy measures is the topic of Section 6.

## 5. FINITE DIFFERENCE SCHEMES

In this section we apply the abstract result from Section 4 to a class of monotone finite difference schemes. To simplify matters we assume that the second order coefficient  $\sigma$  is independent of  $x$ . The general case is much harder and will be addressed in another paper. We also assume  $\nu(dz)$  is a bounded measure with compact support in  $E$  (in which case we may relax slightly the assumptions on the jump vector  $\eta$ ). The general case can always be reduced to this case by applying the truncation procedures discussed in Section 6.

We use the following assumptions in this section:

(A.1')  $\nu$  is a positive Radon measure on  $E$  such that

$$\int_E e^{\ell|z|} \nu(dz) < \infty \quad \text{and} \quad \int_{E \setminus B(0,R)} \nu(dz) = 0$$

for some given  $\ell > 0$  and  $R > 0$ .

(A.2') There exists  $M > 0$  such that for any  $\theta \in \Theta$

$$|\sigma^\theta|_1 + |b^\theta|_1 + |c^\theta|_1 + |f^\theta|_1 \leq M \quad \text{and} \quad |\eta^\theta(\cdot, z)|_1 \leq M e^{\ell|z|}.$$

Since  $\nu(dz)$  is a bounded measure, we may rewrite (1.1) in the form

$$(5.1) \quad \sup_{\theta \in \Theta} \left\{ -\operatorname{tr} \left[ \frac{1}{2} a^\theta D^2 u(x) \right] - \bar{b}^\theta(x) Du(x) + c^\theta(x) u(x) - \mathcal{I}u(x) - f^\theta(x) \right\} = 0,$$

where

$$\begin{aligned} \bar{b}^\theta(x) &= b^\theta(x) - \int_{E_1} \eta^\theta(x, z) \nu(dz), \\ \mathcal{I}u(x) &= \int_E \left( u(x + \eta^\theta(x, z)) - u(x) \right) \nu(dz). \end{aligned}$$

Observe that by (A.1') and (A.2') there holds  $\bar{b}^\theta \in C^{0,1}(\mathbb{R}^N)$ .

Let us now discuss how to discretize the integro-partial differential equation (5.1). We begin with the integro part of this equation. Note that any reasonable approximation of the integral  $\int_E f(z) \nu(dz)$  has the form  $\sum_{\alpha \in \mathcal{P}} f(Z_\alpha) \omega_\alpha$ , where  $\alpha = (\alpha_1, \dots, \alpha_M) \in \mathcal{P} \subset \mathbb{Z}^M$  is a multiindex and  $\{Z_\alpha\}_\alpha$  and  $\{\omega_\alpha\}_\alpha$  denote respectively the nodes and weights specific to each integration method.

To simplify matters we take  $Z_\alpha = h_2 \alpha$ ,  $\alpha \in \mathcal{P}$ , and consider the following quadrature rule:

$$(5.2) \quad Q_{h_2}[f] := \sum_{\alpha \in \mathcal{P}} f(h_2 \alpha) \omega_\alpha \approx \int_E f(z) \nu(dz).$$

By (A.1') we may assume that  $\mathcal{P} \subset \mathbb{Z}^M$  is bounded. Note that if  $\omega_\alpha \geq 0$  for every  $\alpha \in \mathcal{P}$ , then this approximation is obviously monotone. We assume that the following error bound holds for some  $\bar{n} \in \mathbb{N}$ :

$$(5.3) \quad E_{h_2}[\phi] := \left| \int_E \phi(z) \nu(dz) - Q_{h_2}[\phi] \right| \leq C h_2^{\bar{n}} |D^{\bar{n}} \phi|_0,$$

for every sufficiently smooth function  $\phi$ . This bound is satisfied by most commonly used quadrature rules, in particular it is satisfied for all compound rules based on

simple rules that are exact for polynomials of degree  $\bar{n}$ . This class includes the Newton-Cotes and Gauss rules as well as multidimensional product rules based on these rules. We discuss the Newton-Cotes rules in Subsection 5.2. For a general reference on numerical integration we refer to [25].

By discretizing all integrals in (5.1) using the quadrature rule  $Q_{h_2}$  at every point  $x \in \mathbb{R}^N$ , we get the following intermediate approximation of (5.1):

$$(5.4) \quad \sup_{\theta \in \Theta} \left\{ -\operatorname{tr} \left[ \frac{1}{2} a^\theta D^2 u(x) \right] - \bar{b}^{\theta, h_2}(x) Du(x) \right. \\ \left. + c^\theta(x) u(x) - Q_{h_2}[u(x + \eta^\theta(x, \cdot)) - u(x)] - f^\theta(x) \right\} = 0,$$

where we have set

$$\bar{b}^{\theta, h_2}(x) := b^\theta(x) - Q_{h_2}[\eta^\theta(x, \cdot) \mathbf{1}_{B(0,1)}].$$

Note that by assumptions (A.1') and (A.2'), we know that  $|Q_{h_2}[\eta^\theta]|_1$  and  $|\bar{b}^{\theta, h_2}|_1$  are bounded independently of  $h_2$  and  $\theta$ .

Next we discretize the differential equation part of (5.4). For this purpose we use a straightforward monotone finite difference scheme. This scheme is studied in the book [43]. To define the scheme, we need the following finite difference operators:

$$\begin{aligned} \Delta_{x_i}^\pm w(x) &= \pm \frac{1}{h_1} [w(x \pm e_i h_1) - w(x)], \\ \Delta_{x_1}^2 w(x) &= \frac{1}{h_1^2} [w(x + e_1 h_1) - 2w(x) + w(x - e_1 h_1)], \\ \Delta_{x_i x_j}^+ w(x) &= \frac{1}{2h_1^2} [2w(x) + w(x + e_i h_1 + e_j h_1) + w(x - e_i h_1 - e_j h_1)] \\ &\quad - \frac{1}{2h_1^2} [w(x + e_i h_1) + w(x - e_i h_1) + w(x + e_j h_1) + w(x - e_j h_1)], \\ \Delta_{x_i x_j}^- w(x) &= \frac{1}{2h_1^2} [w(x + e_i h_1) + w(x - e_i h_1) + w(x + e_j h_1) + w(x - e_j h_1)] \\ &\quad - \frac{1}{2h_1^2} [2w(x) + w(x + e_i h_1 - e_j h_1) + w(x - e_i h_1 + e_j h_1)]. \end{aligned}$$

Using the notation  $z^+ := \max(z, 0)$  and  $z^- := (-z)^+$  (so that  $z = z^+ - z^-$  and  $|z| = z^+ + z^-$ ) for any real number  $z$ , we can then write the full scheme as

$$(5.5) \quad \hat{H}(h_2, x, u_h(x), \Delta_{x_i}^\pm u_h(x), \Delta_{x_i}^2 u_h(x), \Delta_{x_i x_j}^\pm u_h(x), u_h(\cdot)) = 0,$$

where

$$\begin{aligned} &\hat{H}(h_2, x, r, p_i^\pm, A_{ii}, A_{ij}^\pm, u(\cdot)) \\ &:= \sup_{\theta \in \Theta} \left\{ \sum_{i=1}^N \left[ -\frac{a_{ii}^\theta}{2} A_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^N \left( -\frac{(a_{ij}^\theta)^+}{2} A_{ij}^+ + \frac{(a_{ij}^\theta)^-}{2} A_{ij}^- \right) \right. \right. \\ &\quad \left. \left. - (\bar{b}_i^{\theta, h_2})^+(x) p_i^+ + (\bar{b}_i^{\theta, h_2})^-(x) p_i^- \right] \right. \\ &\quad \left. - Q_{h_2}[u(x + \eta^\theta(x, z)) - u(x)] + c^\theta(x) r - f^\theta(x) \right\}. \end{aligned}$$

Now we assume that (A.1'), (A.2'), and (A.3) holds, and  $h_1, h_2 < 1$ . In addition, we use the following assumptions:

$$(B.1) \quad a^\theta := \sigma^\theta \sigma^{\theta^\top} \quad \text{is a constant matrix.}$$

$$(B.2) \quad a_{ii}^\theta - \sum_{\substack{j=1 \\ j \neq i}}^N |a_{ij}^\theta| \geq 0, \quad \theta \in \Theta, \quad i = 1, \dots, N \quad \text{and} \quad \omega_\alpha \geq 0, \quad \alpha \in \mathcal{P}.$$

$$(B.3) \quad \sum_{i=1}^N \left[ a_{ii}^\theta - \sum_{\substack{j=1 \\ j \neq i}}^N |a_{ij}^\theta| + |\bar{b}_i^{\theta, h_2}|_0 + \sum_{\alpha \in \mathcal{P}} \omega_\alpha \right] \leq 1, \quad \theta \in \Theta, \quad h_2 < 1.$$

$$(B.4) \quad |D_z^{\bar{n}} \eta^\theta(x, \cdot) \mathbf{1}_{B(0, R)}|_0 \leq C, \quad x \in \mathbb{R}^N, \quad \theta \in \Theta, \quad \text{where } \bar{n}, R \text{ are defined in (5.3), (A.1')} \text{ respectively.}$$

Assumption (B.1) is used to simplify the proof of condition (D) in Section 4 for the scheme (5.5). The general case is much more involved and will be considered elsewhere. Assumption (B.2) implies that the scheme is monotone and contains the standard requirement that  $a$  should be diagonally dominant. Assumption (B.3) will always hold after a proper rescaling of the equation. Finally, the regularity assumption (B.4) is necessary if we want to use the error bound (5.3) when the quadrature rule (5.2) is applied to  $\eta^\theta$  and  $u(x + \eta^\theta)$ . This assumption can be dropped if we replace  $\eta^\theta$  in (1.1) by a smooth approximation and then keep proper track of the error that we make in the process. However, since this additional step complicates the presentation, we skip it and employ instead condition (B.4).

**Theorem 5.1.** *Assume that (A.1'), (A.2'), (A.3), (B.1) - (B.4), and (5.3) hold. Assume also that  $\lambda > \bar{\lambda}_0$  (cf. Remark 4.5) and that  $h_1, h_2 < 1$  are sufficiently small. Let  $u$  be the viscosity solution of (5.1) and  $u_h$  be the solution of the finite difference scheme (5.5). Then*

$$|u - u_h|_0 \leq C(h_1^{1/2} + h_2).$$

Before proving this theorem we will give several remarks and an example.

**5.1. Remarks.** The result  $Ch_1^{1/2}$  for the pure differential part is consistent with previous results by Krylov and Barles-Jakobsen [40, 41, 8]. We refer to [35] where the rate 1/2 is derived in the pure PDE case. Note that by increasing the order of the quadrature rule (i.e., increasing  $\bar{n}$ ) one does not improve the corresponding order in the error bound; the result is always  $Ch_2$ . This fact and the reduction of rate observed in this result is a consequence of the lack of regularity of  $u$  (the viscosity solution  $u$  is here merely Lipschitz continuous). In fact when the Lévy measure is bounded and the solution is Lipschitz continuous, we can always take  $\bar{n} = 1$  in (5.3) for any monotone quadrature formula that integrates constants exactly (this includes most commonly used quadrature rules). To illustrate this claim, let us check it for the Riemann sum approximation, i.e., (5.2) with weights  $\omega_\alpha = \nu(h_2\alpha + [0, h_2]^M)$ . Indeed, we have

$$\begin{aligned} & \sum_{\alpha \in \mathcal{P}} f(h_2\alpha) \omega_\alpha - \int_E f(z) \nu(dz) \\ &= \sum_{\alpha \in \mathcal{P}} \int_{h_2\alpha + [0, h_2]^M} (f(h_2\alpha) - f(z)) \nu(dz) \\ &= \sum_{\alpha \in \mathcal{P}} \int_{h_2\alpha + [0, h_2]^M} \int_0^1 Df((1-t)h_2\alpha + tz)(z - h_2\alpha) dt \nu(dz), \end{aligned}$$

which implies the following error bound (cf. (5.3)):

$$E_{h_2}[f] \leq h_2 \int_E |Df(z)| \nu(dz) \leq |\nu(E)| |Df|_0 h_2.$$

Combining this with Theorem 5.1 yields



**Corollary 5.2.** *Suppose all the assumptions in Theorem 5.1 hold with  $\bar{n} = 1$  and that the quadrature rule in (5.5) is based on a simple quadrature rule that integrates constants exactly. Then*

$$|u - u_h|_0 \leq C(h_1^{1/2} + h_2).$$

**5.2. The Newton-Cotes formulas.** We now discuss the Newton-Cotes formulas, see, e.g., [25]. These formulas are the simplest one-dimensional quadrature formulas and include the trapezoidal and Simpson's rules. The simple Newton-Cotes formulas are obtained by (i) dividing the interval of integration  $[a, b]$  into  $\bar{n}$  equally sized subintervals, yielding  $\bar{n} + 1$  equidistributed nodes when the endpoints are included, (ii) interpolating the integrand using Lagrange interpolation and the nodes from (i), and (iii) integrating the Lagrange interpolant (which is a polynomial). This leads to the formula

$$(5.6) \quad Q_{sNC, \bar{n}}[f; a, b] := \frac{b-a}{\bar{n}} \sum_{i=0}^{\bar{n}} f(z_i) \omega_i \approx \int_a^b f(z) \nu(dz),$$

where the subscript "sNC" stands for "simple Newton-Cotes" and

$$\omega_i = (-1)^{\bar{n}-i} \frac{\bar{n}}{b-a} \int_0^{\bar{n}} \frac{t(t-1) \dots (t-\bar{n})}{(t-i)!(\bar{n}-i)!} \nu(d\phi(t)) \quad \text{where} \quad \phi(t) = a + (b-a)t.$$

When  $\nu(dz) = dz$  the first few rules are

$$\begin{aligned} Q_{sNC,1}[f; a, b] &= h \frac{f_0 + f_1}{2} \quad (\text{trapezoidal rule}), \\ Q_{sNC,2}[f; a, b] &= \frac{h}{6} (f_0 + 4f_1 + f_2) \quad (\text{Simpson's rule}), \\ Q_{sNC,3}[f; a, b] &= \frac{3h}{8} (f_0 + 3f_1 + 3f_2 + f_3), \end{aligned}$$

where  $h = (b-a)/\bar{n}$  and  $f_i = f(a + hi)$ ,  $i = 0, \dots, \bar{n}$ .

For  $\bar{n} \leq 8$  the simple Newton-Cotes formulas are monotone. In practice it is the compound formulas  $Q_{NC, \bar{n}}[f; a, b]$  that are used, that is, we subdivide  $[a, b]$  into  $N_s$  equally sized intervals and apply the simple quadrature rule on each interval:

$$(5.7) \quad Q_{NC, \bar{n}}[f; a, b] := \sum_{i=0}^{N_s-1} Q_{sNC, \bar{n}}[f; y_i, y_{i+1}] \approx \int_a^b f(z) \nu(dz),$$

where  $y_i = a + i \frac{b-a}{N_s}$ . In the compound rules there are  $N_s \bar{n} + 1$  nodes and the step size is  $h = \frac{b-a}{N_s \bar{n}}$ . If the integrand  $f$  is sufficiently smooth, the error can be bounded as follows:

$$E_{NC, \bar{n}}[f] := \left| \int_a^b f(z) dz - Q_{NC, \bar{n}}[f] \right| \leq \begin{cases} C_{\bar{n}} h^{\bar{n}} |D^{\bar{n}} f|_0, & \text{for } \bar{n} \text{ odd,} \\ C_{\bar{n}} h^{\bar{n}+1} |D^{\bar{n}+1} f|_0, & \text{for } \bar{n} \text{ even,} \end{cases}$$

where  $C_{\bar{n}}$  is a constant. An immediate consequence of Theorem 5.1 is

**Corollary 5.3.** *Suppose all the assumptions in Theorem 5.1 hold, and let  $M = 1$  (so  $z \in \mathbb{R}$ ),  $\bar{n} \leq 8$ , and  $Q_{NC, \bar{n}}$  be the quadrature rule in (5.5). Then*

$$|u - u_h|_0 \leq C(h_1^{1/2} + h_2).$$

*Remark 5.4.* In higher dimensions ( $M > 1$ ) we can apply one-dimensional Newton-Cotes rules dimension-by-dimension. If the grid is a subset of  $h_2 \mathbb{Z}^M$ , then the corresponding approximation error for these product rules is of the same order in  $h_2$  as in the one-dimensional case. Moreover, Corollary 5.3 holds.

**5.3. Proof of Theorem 5.1.** Theorem 5.1 will be a consequence of Theorem 4.4 once we have verified conditions (C.1) - (C.5) and (D). To this end, we will rewrite the scheme (5.5) as a discrete dynamic programming principle.

For  $x \in \mathbb{R}^N$  and  $\theta \in \Theta$ , define the following “transition probabilities”:

$$\begin{aligned} p^\theta(x, x) &= 1 - \sum_{i=1}^N \left[ a_{ii}^\theta - \sum_{\substack{j=1 \\ j \neq i}}^N |a_{ij}^\theta| - h_1 |\bar{b}_i^{\theta, h_2}(x)| \right] - h_1^2 \sum_{\substack{\alpha \in \mathcal{P} \\ \alpha \neq 0}} \omega_\alpha(x), \\ p^\theta(x, x \pm e_i h_1) &= \frac{a_{ii}^\theta}{2} - \sum_{\substack{j=1 \\ j \neq i}}^N \frac{|a_{ij}^\theta|}{2} + h_1 (\bar{b}_i^{\theta, h_2})^\pm(x), \quad i = 1, \dots, N, \\ p^\theta(x, x + e_i h_1 \pm e_j h_1) &= \frac{(a_{ij}^\theta)^\pm}{2}, \quad i, j = 1, \dots, N, \quad i \neq j, \\ p^\theta(x, x - e_i h_1 \pm e_j h_1) &= \frac{(a_{ij}^\theta)^\mp}{2}, \quad i, j = 1, \dots, N, \quad i \neq j, \\ p^\theta(x, x + \eta^\theta(x, \alpha h_2)) &= h_1^2 \omega_\alpha, \quad \alpha \in \mathcal{P}, \quad \alpha \neq 0, \end{aligned}$$

and  $p^\theta(x, y) = 0$  for every other  $y$ .

If  $0 < h_1 < 1$  then (B.1) - (B.3) implies for all  $x, y$ , and  $\theta$  that

$$0 \leq p^\theta(x, y) \leq 1 \quad \text{and} \quad \sum_{z \in h_1 \mathbb{Z}^N \cup \eta^\theta(x, h_2 \mathcal{P})} p^\theta(x, x + z) = 1.$$

Tedious but straightforward computations will reveal that we can rewrite the scheme (5.5) in the form

$$(5.8) \quad \begin{aligned} &u_h(x) \\ &= \sup_{\theta \in \Theta} \left\{ \frac{1}{1 + c^\theta(x) h_1^2} \left[ \sum_{z \in h_1 \mathbb{Z}^N \cup \eta^\theta(x, h_2 \mathcal{P})} p^\theta(x, x + z) u_h(x + z) + h_1^2 f^\theta(x) \right] \right\}. \end{aligned}$$

We use this expression to define the functional  $S$  in (1.4). For each triplet  $(x, r, \phi) \in (\mathbb{R}^N, \mathbb{R}, C_b(\mathbb{R}^N))$ , we define  $S$  by

$$(5.9) \quad \begin{aligned} &S(h, x, r, [\phi]_x) \\ &:= \sup_{\theta \in \Theta} \left\{ -\frac{1}{h_1^2} \left[ \sum_{z \in h_1 \mathbb{Z}^N \cup \eta^\theta(x, h_2 \mathcal{P})} p^\theta(x, x + z) [\phi]_x(z) - r \right] + c^\theta(x) r - f^\theta(x) \right\}, \end{aligned}$$

where we have set  $[\phi]_x(z) := \phi(x + z)$ . It is easy to see that (1.4), with  $S$  as defined in (5.9), is equivalent to (5.8).

By our assumptions and definitions of transition probabilities  $p^\theta(x, y)$ , this scheme is obviously monotone. It is convex since it is defined by taking the supremum of a family of linear expressions. Moreover, if  $\phi$  is a smooth function, then an argument based on Taylor expansions yields

$$(5.10) \quad \begin{aligned} &|H(x, v, Dv, D^2 v, v(\cdot)) - S(h, x, v(x), [v]_x)| \\ &\leq C \sup_{\theta \in \Theta} \left( |D^2 v|_0 h_1 + |D^4 v|_0 h_1^2 + |E_{h_2}[v(\cdot + \eta^\theta) - v]|_0 + [v]_1 |E_{h_2}[\eta^\theta]|_0 \right) \\ &\leq C \sup_{\theta \in \Theta} \left( |D^2 v|_0 h_1 + |D^4 v|_0 h_1^2 \right. \\ &\quad \left. + (|D_z^{\bar{n}} v(\cdot + \eta^\theta) \mathbf{1}_{B(0, R)}|_0 + [v]_1 |D_z^{\bar{n}} \eta^\theta \mathbf{1}_{B(0, R)}|_0) h^{\bar{n}} \right). \end{aligned}$$

Summarizing, we have proved

**Lemma 5.5.** *Suppose the hypotheses of Theorem 5.1 hold. Then the scheme (5.5) satisfies conditions (C.1) - (C.5), where (C.3) takes the form (5.10).*

It remains to prove that condition (D) holds. From now on we make the assumption that  $c(x) = \lambda$  for all  $x$  (the general case is similar but algebraically more involved). Let  $T_h$  be the operator defined by

$$T_h v(x) := \frac{1}{1 + \lambda h_1^2} \sup_{\theta \in \Theta} \left\{ \sum_{z \in h_1 \mathbb{Z}^N \cup \eta^\theta(x, h_2 \mathcal{P})} p^\theta(x, x+z) v(x+z) + h_1^2 f^\theta(x) \right\}.$$

This operator is a contraction on  $C_b(\mathbb{R}^N)$ , since

$$\begin{aligned} & |T_h u(x) - T_h v(x)| \\ & \leq \frac{1}{1 + \lambda h_1^2} \sup_{\theta \in \Theta} \sum_{z \in h_1 \mathbb{Z}^N \cup \eta^\theta(x, h_2 \mathcal{P})} p^\theta(x, x+z) |u(x+z) - v(x+z)| \\ & \leq \frac{1}{1 + \lambda h_1^2} \|u - v\|_0. \end{aligned}$$

Since (5.8) reads  $u_h = T_h u_h$ , Banach's fixed point theorem yields existence and uniqueness of a solution  $u_h \in C_b(\mathbb{R}^N)$  to (5.8).

Now we proceed to show that this solution is Lipschitz continuous. Assume that  $\phi \in C^{0,1}(\mathbb{R}^N)$ . For any  $x, y \in \mathbb{R}^N$ ,

$$\begin{aligned} & T_h \phi(x) - T_h \phi(y) \\ & = \frac{1}{1 + \lambda h_1^2} \sup_{\theta \in \Theta} \left\{ \sum_{z \in h_1 \mathbb{Z}^N \cup \eta^\theta(x, h_2 \mathcal{P})} \left[ p^\theta(x, x+z) (\phi(x+z) - \phi(y+z)) \right. \right. \\ & \quad \left. \left. + (p^\theta(x, x+z) - p^\theta(y, y+z)) \phi(y+z) \right] + h_1^2 [f^\theta(x) - f^\theta(y)] \right. \\ & \quad \left. + \left[ \sum_{z \in h_1 \mathbb{Z}^N \cup \eta^\theta(x, h_2 \mathcal{P})} - \sum_{z \in h_1 \mathbb{Z}^N \cup \eta^\theta(y, h_2 \mathcal{P})} \right] p^\theta(y, y+z) \phi(y+z) \right\}. \end{aligned}$$

Clearly, we have

$$\sum_{z \in h_1 \mathbb{Z}^N \cup \eta^\theta(x, h_2 \mathcal{P})} p^\theta(x, x+z) (\phi(x+z) - \phi(y+z)) \leq [\phi]_1 |x - y|.$$

Using (A.1'), (A.2'), (5.3), assuming  $h_2$  to be sufficiently small, and taking into account the definitions of  $p^\theta(x, x+z)$ ,  $\bar{b}^{\theta, h_2}$ , we find

$$\begin{aligned} & \sum_{z \in h_1 \mathbb{Z}^N \cup \eta^\theta(x, h_2 \mathcal{P})} (p^\theta(x, x+z) - p^\theta(y, y+z)) \phi(y+z) \\ & = h_1^2 \sum_{i=1}^N \left[ \left( \bar{b}_i^{\theta, h_2+}(x) - \bar{b}_i^{\theta, h_2+}(y) \right) \Delta_{x_i}^+ \phi(y) - \left( \bar{b}_i^{\theta, h_2-}(x) - \bar{b}_i^{\theta, h_2-}(y) \right) \Delta_{x_i}^- \phi(y) \right] \\ & \leq h_1^2 [\phi]_1 2N ([b^\theta]_1 + [Q_{h_2}[\eta^\theta]]_1) |x - y| \\ & \leq h_1^2 [\phi]_1 2NM \left( 1 + 2 \int_E e^{\ell|z|} \nu(dz) \right) |x - y|. \end{aligned}$$

Due to (A.1'), (A.2'), and (5.3), we also have

$$\begin{aligned} & \left[ \sum_{z \in h_1 \mathbb{Z}^N \cup \eta^\theta(x, h_2 \mathcal{P})} - \sum_{z \in h_1 \mathbb{Z}^N \cup \eta^\theta(y, h_2 \mathcal{P})} \right] p^\theta(y, y+z) \phi(y+z) \\ &= h_1^2 \sum_{\alpha \in \mathcal{P}} \omega_\alpha \left( \phi(y + \eta^\theta(x, \alpha h_2)) - \phi(y + \eta^\theta(y, \alpha h_2)) \right) \\ &\leq h_1^2 [\phi]_1 M 2 \int_E e^{\ell|z|} \nu(dz) |x - y|. \end{aligned}$$

Hence we conclude that

$$T_h \phi(x) - T_h \phi(y) \leq \frac{1}{1 + \lambda h_1^2} \left[ (1 + \bar{\lambda}_0 h_1^2) [\phi]_1 + h_1^2 \sup_{\theta \in \Theta} [f^\theta]_1 \right] |x - y|,$$

where

$$\bar{\lambda}_0 := M(2N + 1) \left( 1 + 2 \int_E e^{\ell|z|} \nu(dz) \right).$$

By interchanging  $x$  and  $y$  we have now proved that

$$[T_h \phi]_1 \leq \frac{1 + \bar{\lambda}_0 h_1^2}{1 + \lambda h_1^2} [\phi]_1 + \frac{h_1^2}{1 + \lambda h_1^2} \sup_{\theta \in \Theta} [f^\theta]_1 =: A[\phi]_1 + B.$$

The constant  $\bar{\lambda}_0$  coincides with the one mentioned in Remark 4.5 for this case. By assumption  $\lambda > \bar{\lambda}_0$  and hence  $A < 1$ . By the fixpoint theorem mentioned above, for any  $\phi \in C^{0,1}(\mathbb{R}^N) \subset C_b(\mathbb{R}^N)$ ,

$$|T_h^n \phi - u_h|_0 \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

where  $u_h$  is the unique in  $C_b(\mathbb{R}^N)$  solution of (5.8). Furthermore

$$[u_h]_1 = \lim_{n \rightarrow \infty} [T_h^n \phi]_1 \leq \lim_{n \rightarrow \infty} \left\{ A^n [\phi]_1 + \sum_{i=0}^{n-1} A^i B \right\} = \frac{B}{1 - A} = \frac{1}{\lambda - \bar{\lambda}_0} \sup_{\theta \in \Theta} [f^\theta]_1.$$

We have proved:

**Lemma 5.6.** *Suppose the hypotheses of Theorem 5.1 hold. Then there exists a unique solution  $u_h \in C^{0,1}(\mathbb{R}^N)$  to the scheme (5.5).*

Next we consider the perturbed scheme (1.6), with  $S$  as defined in (5.9). Similarly to what we did above, the perturbed scheme can be written as

$$(5.11) \quad \begin{aligned} & u_h^\varepsilon(x) \\ &= \frac{1}{1 + \lambda h_1^2} \sup_{\substack{\theta \in \Theta \\ |e| \leq \varepsilon}} \sum_{z \in h_1 \mathbb{Z}^N \cup \eta^\theta(x+e, h_2 \mathcal{P})} \left\{ p_e^\theta(x, x+z) u_h^\varepsilon(x+z) + h_1^2 f^\theta(x+e) \right\}, \end{aligned}$$

where the coefficients  $p_e^\theta(x, x+z)$  are obtained from  $p^\theta(x, x+z)$  by replacing  $(\bar{b})^\pm(x)$  with  $(\bar{b})^\pm(x+e)$ . Existence and uniqueness of a  $C^{0,1}(\mathbb{R}^N)$  solution to the perturbed scheme follows from Lemma 5.6 once we redefine  $\theta$  and  $\Theta$  to  $(\theta, e)$  and  $\Theta \times \{|e| \leq \varepsilon\}$ , respectively, and rewrite (5.11) in the form (5.8) (or (5.5)). Note that the constant  $\bar{\lambda}_0$  remains the same.

It remains to prove that  $|u_h - u_h^\varepsilon|_0 \leq C\varepsilon$ . It is easy to see that  $u_h^\varepsilon$ , the solution of equation (5.11), is a subsolution of equation (5.8). By the comparison principle (cf. Lemma 4.1),  $u_h^\varepsilon \leq u_h$ . To get the other bound, let  $T_h^\varepsilon u_h^\varepsilon$  denote the right-hand side of (5.11). A computation similar to the one used above to prove the Lipschitz

continuity of  $u_h$  then leads to

$$\begin{aligned} & T_h u_h(x) - T_h^\varepsilon u_h^\varepsilon(x) \\ & \leq \frac{1}{1 + \lambda h_1^2} \left\{ |u_h - u_h^\varepsilon|_0 + C h_1^2 \sup_{\substack{\theta \in \Theta \\ |e| \leq \varepsilon}} \left\{ [u_h^\varepsilon]_1 |\bar{b}^{\theta, h_2} - \bar{b}^{\theta, h_2}(\cdot + e)|_0 \right. \right. \\ & \quad \left. \left. + [u_h^\varepsilon]_1 |Q_{h_2}[\eta^\theta(\cdot, z) - \eta^\theta(\cdot + e, z)]|_0 + |f^\theta - f^\theta(\cdot + e)|_0 \right\} \right\}. \end{aligned}$$

By Lemma 5.6,  $|u_h^\varepsilon|_1$  is bounded independently of  $\varepsilon$  and  $h$ , so by (A.1') and (A.2') we have

$$u_h(x) - u_h^\varepsilon(x) = T_h u_h(x) - T_h^\varepsilon u_h^\varepsilon(x) \leq \frac{1}{1 + \lambda h_1^2} \left[ |u_h - u_h^\varepsilon|_0 + C \varepsilon h_1^2 \right],$$

which immediately implies  $u_h(x) - u_h^\varepsilon(x) \leq C\varepsilon$ . Summarizing our findings, we have just proved

**Lemma 5.7.** *Suppose the hypotheses of Theorem 5.1 hold. Then the scheme (5.5) satisfies condition (D).*

This completes the proof of Theorem 5.1.

## 6. GENERAL LÉVY MEASURES

The finite difference schemes for (1.1) in Section 5 were constructed under the assumption that the Lévy measure was bounded and also compactly supported. In this section we show how one can reduce problems with general unbounded and non-compactly supported Lévy measures to this case. Once such a reduction process has been carried out, we can use for example the finite difference schemes from Section 5. The reduction process is based on appropriate truncations (through two scales) of the Lévy measure, and we provide precise estimates for the error induced by the truncations. Hence, the desired "level of accuracy" related to the truncations can be determined a priori before any numerical discretization is applied. We should emphasize that in this section  $\sigma$  is not assumed to be independent of  $x$ !

To be more precise, let us introduce the "two-scales" truncated Lévy measure

$$\nu_{r,R}(dz) := \mathbf{1}_{r < |z| < R} \nu(dz),$$

for some (small)  $r \in (0, 1)$  and for some (large)  $R \in (1, \infty)$ . Clearly,  $\nu_{r,R}(dz)$  is a bounded and compactly supported measure. For  $\phi \in C^\infty(\mathbb{R}^N)$ , let us express the integro operator  $\mathcal{J}^\theta$  in (1.1) as follows:

$$(6.1) \quad \mathcal{J}^\theta \phi(x) = \mathcal{J}_r^\theta \phi(x) + \mathcal{J}_{r,R}^\theta \phi(x) + \mathcal{J}^{\theta,R} \phi(x),$$

where

$$\begin{aligned} \mathcal{J}_r^\theta \phi(x) &:= \int_{0 < |z| \leq r} \left[ \phi(x + \eta^\theta(x, z)) - \phi(x) - \eta^\theta(x, z) D\phi(x) \right] \nu(dz), \\ \mathcal{J}_{r,R}^\theta \phi(x) &:= \int_E \left[ u(x + \eta^\theta(x, z)) - u(x) - \mathbf{1}_{|z| < 1} \eta^\theta(x, z) Du(x) \right] \nu_{r,R}(dz), \\ \mathcal{J}^{\theta,R} \phi(x) &:= \int_{|z| \geq R} \left[ \phi(x + \eta^\theta(x, z)) - \phi(x) \right] \nu(dz). \end{aligned}$$

We want to take a closer look at the decomposition (6.1). To this end, we divide the remaining part of this section into two parts.

**6.1. Treatment of the operators  $\mathcal{J}_r^\theta$  and  $\mathcal{J}^{\theta,R}$ .** To reduce  $\nu(dz)$  to a compactly supported measure, we simply throw away the term  $\mathcal{J}^{\theta,R}\phi(x)$  in (6.1), and observe that by (A.1) and (A.2)

$$|\mathcal{J}^{\theta,R}\phi(x)| \leq 2|\phi|_0 \int_{|z| \geq R} \nu(dz) \leq C|\phi|_0 e^{-\ell R}.$$

Next we discuss how to reduce  $\nu(dz)$  to a bounded measure. Roughly speaking, the idea is to get rid of the singularity of the Lévy measure  $\nu(dz)$  by suppressing the "small jumps", i.e., by throwing away the operator  $\mathcal{J}_r^\theta$  in (6.1). It does not, however, result in a good approximation if we completely suppress the small jumps. Indeed, the small jumps are important in most of the Lévy models found in the literature. Therefore, to account for the small jumps (and thus the infinite activity of the Lévy process), we follow [29, 30] and replace  $\mathcal{J}_r^\theta\phi$  in (6.1) by a diffusion term:

$$\mathcal{J}_r^\theta\phi \approx \text{tr}[a_r^\theta D^2\phi],$$

where the diffusion matrix  $a_r^\theta$  needs to be specified.

There are (infinitely) many choices for the diffusion matrix  $a_r^\theta$ , some yielding more accurate results than others. We quantify the accuracy of some choices of  $a_r^\theta$  below. Define

$$\omega^\theta(x, r) := \frac{1}{2} \int_{0 < |z| \leq r} \eta^\theta(x, z) \eta^\theta(x, z)^\top \nu(dz),$$

and consider the three choices

$$\begin{aligned} a_{1,r}^\theta &:= \omega^\theta(x, r), \\ a_{2,r}^\theta &:= \lim_{\rho \rightarrow \infty} \frac{1}{|B(0, \rho)|} \int_{B(0, \rho)} \omega^\theta(x, r) dx, \\ a_{3,r}^\theta &:= 0. \end{aligned}$$

Notice that  $a_2^\theta$  and  $a_3^\theta$  are independent of  $x$ . Also,  $a_{2,r}^\theta$  is meaningful since  $\eta^\theta$  is bounded in  $x$ , cf. (A.2).

**Lemma 6.1.** *Supppose (A.1) and (A.2) hold, and let  $\phi \in C^\infty(\mathbb{R}^N)$ . Then the following error estimates hold:*

$$\begin{aligned} |\mathcal{J}_r^\theta\phi - \text{tr}[a_{1,r}^\theta D^2\phi]|_0 &\leq |D^3\phi|_0 r \sup_{\theta \in \Theta} |\omega^\theta(\cdot, r)|_0, \\ |\mathcal{J}_r^\theta\phi - \text{tr}[a_{2,r}^\theta D^2\phi]|_0 &\leq |D^2\phi|_0 \sup_{\theta \in \Theta} |\omega^\theta(\cdot, r) - a_{2,r}^\theta|_0 + |D^3\phi|_0 r \sup_{\theta \in \Theta} |\omega^\theta(\cdot, r)|_0, \\ |\mathcal{J}_r^\theta\phi - \text{tr}[a_{3,r}^\theta D^2\phi]|_0 &\leq |D^2\phi|_0 \sup_{\theta \in \Theta} |\omega^\theta(\cdot, r)|_0. \end{aligned}$$

*Proof.* By Taylor expanding  $\phi(x + \eta^\theta(x, z))$  we find

$$\begin{aligned} \mathcal{J}_r^\theta\phi(x) &= \frac{1}{2} \text{tr} \left[ D^2\phi(x) \int_{0 < |z| \leq r} \eta^\theta(x, z) \eta^\theta(x, z)^\top \nu(dz) \right] \\ &\quad + \frac{1}{2} \text{tr} \left[ \int_{0 < |z| \leq r} \left( D^2\phi(\zeta_{x,z}) - D^2\phi(x) \right) \eta^\theta(x, z) \eta^\theta(x, z)^\top \nu(dz) \right], \end{aligned}$$

where  $\zeta_{x,z}$  belongs to the straight line segment between  $x$  and  $x + \eta^\theta(x, z)$ . Using this, and conditions (A.1) and (A.2), one can verify the assertions of the lemma.  $\square$

We have here three "diffusion approximations"  $\mathcal{J}_r^\theta\phi \approx \text{tr}[a_{l,r}^\theta D^2\phi]$ ,  $l = 1, 2, 3$ . According to Lemma 6.1, the first choice has the highest order of convergence as  $r \rightarrow 0$ . The second and third choices have the same order of convergence, but the error is smaller for the second alternative when  $r$  is small.

**6.2. Error estimates.** In view of what we have said up to now, it seems that the following problems may serve as “good” approximations to (1.1), at least when  $r$  is small and  $R$  is large:

$$(6.2) \quad H_{i,r,R}(x, u(x), Du(x), D^2u(x), u(\cdot)) = 0 \quad \text{in } \mathbb{R}^N, \quad i = 1, 2, 3,$$

where, for any  $(x, r, p, X) \in \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}^N \times \mathbb{S}^N$  and any locally bounded function  $u$ , the Hamiltonian  $H_{i,r,R} = H_{i,r,R}(x, r, p, X, u(\cdot))$  is defined by

$$(6.3) \quad \begin{aligned} & H_{i,r,R}(x, r, p, X, u(\cdot)) \\ &= \sup_{\theta \in \Theta} \left\{ -\frac{1}{2} \text{tr} \left[ \left( a^\theta(x) + a_{i,r}^\theta(x) \right) X \right] \right. \\ & \quad \left. - b^\theta(x) \cdot p - \mathcal{J}_{r,R}^\theta u(x) + c^\theta(x)r - f^\theta(x) \right\}. \end{aligned}$$

We reiterate here that the measure  $\nu_{r,R}(dz)$  is bounded and has compact support.

To get an explicit estimate of the truncation error, that is, the error committed when replacing (1.1) by (6.2), we need stronger assumptions on the Lévy measure. In particular, we need to be more precise about the behavior of the Lévy measure near the origin:

**(A.1'')** Assumption (A.1) holds and in addition there is a density  $m : E \rightarrow [0, \infty)$ , such that  $\nu(dz) = m(z)dz$  and

$$m(z) \leq C \frac{1}{|z|^{N+\alpha}} \quad \text{for } z \in E_1.$$

We emphasize that the conditions in (A.1'') are satisfied by all Lévy processes used in the literature to model financial markets, see [21] and Section 2. If (A.1'') holds, in addition to (A.2), then we can have explicit bounds on the modulus  $\omega^\theta(x, r)$  introduced in the previous subsection. Indeed, turning to polar coordinates  $\rho = |z|$ ,  $dz = C\rho^{N-1}d\rho$ , we obtain

$$|\omega^\theta(x, r)| \leq C \int_{0 < |z| < r} |z|^2 \frac{1}{|z|^{N+\alpha}} dz = Cr^{2-\alpha}.$$

We are now in a position to derive an explicit estimate for the error committed by replacing (1.1) by (6.2).

**Theorem 6.2.** *Assume (A.1''), (A.2), (A.3) hold and let  $u \in C^{0,1}(\mathbb{R}^N)$  be the solution of (1.1) and  $u_{i,r,R} \in C^{0,1}(\mathbb{R}^N)$  be the viscosity solution of (6.2) for  $i = 1, 2, 3$ . Then for  $r > 0$  small enough and  $R$  large enough we have*

$$\begin{aligned} |u - u_{1,r,R}|_0 &\leq C_1(r^{1-\alpha/3} + e^{-\ell R}), \\ |u - u_{2,r,R}|_0 &\leq C_2(r^{1-\alpha/2} + e^{-\ell R}), \\ |u - u_{3,r,R}|_0 &\leq C_3(r^{1-\alpha/2} + e^{-\ell R}), \end{aligned}$$

for some constants  $C_1, C_2, C_3$  that are independent of  $r, R$ , and  $C_2 \leq C_3$ .

**Remark 6.3.** The approximate problem (6.2) with  $i = 1$  gives the smallest error, at least asymptotically as  $r \rightarrow 0$  and  $R \rightarrow \infty$ . Estimates in the same spirit have been obtained by Cont and Voltchkova [22] for a one-dimensional jump-diffusion Black and Scholes equation. This equation is a constant coefficients linear parabolic equation with smooth solutions, and this smoothness is also reflected in better error estimates. We should mention that stochastic arguments are used in [22], as opposed to our viscosity solution arguments yielding estimates that hold for fully nonlinear degenerate integro-partial differential equations for which classical solutions do not exist.

*Proof.* We only prove the lower bound and only for the case  $i = 1$ . The other cases and the upper bounds follow along the same lines. Moreover, since the arguments resemble closely those used in the proof of Theorem 4.4, we are rather brief. To simplify the notation we replace  $u_{i,r,R}$  by  $\bar{u}_i$ ,  $i = 1, 2, 3$ , in the proof below.

The goal is to derive a lower bound on  $u - \bar{u}_1$ . As in Section 3, we can find smooth subsolutions  $\bar{u}_{1,\varepsilon}$  of (6.2), where  $\bar{u}_{1,\varepsilon} = \rho_\varepsilon * u_1^\varepsilon$  and  $u_1^\varepsilon$  is the  $C^{0,1}(\mathbb{R}^N)$  viscosity solution of

$$\begin{aligned} \sup_{\substack{\theta \in \Theta \\ |e| \leq \varepsilon}} \Big\{ & -\frac{1}{2} \operatorname{tr} \left[ \left( a^\theta(x+e) + a_{1,r}^\theta(x+e) \right) D^2 u \right] \\ & - b^\theta(x+e) Du + c^\theta(x+e)u - f^\theta(x+e) \\ & - \int_E \left[ u(x + \eta^\theta(x+e, z)) - u - \mathbf{1}_{|z| < 1} \eta^\theta(x+e, z) Du \right] \nu_{r,R}(dz) \Big\} = 0. \end{aligned}$$

Existence, uniqueness, and regularity properties of  $u_1^\varepsilon$  follow from Theorem 7.1, and the fact that  $\bar{u}_{1,\varepsilon}$  is a subsolution of (6.2) is a consequence of Theorem 7.4 (cf. Section 7). By properties of mollifiers it then follows that

$$\bar{u}_{1,\varepsilon} \in C^\infty(\mathbb{R}^N), \quad |\bar{u}_{1,\varepsilon}|_{0,1} \leq C, \quad |D^i \bar{u}_{1,\varepsilon}|_0 \leq C\varepsilon^{1-i}.$$

According to the calculations we did at the beginning of this section, we have

$$\begin{aligned} & H(x, \bar{u}_{1,\varepsilon}, D\bar{u}_{1,\varepsilon}, D^2\bar{u}_{1,\varepsilon}, \bar{u}_{1,\varepsilon}(\cdot)) - H_{1,r,R}(x, \bar{u}_{1,\varepsilon}, D\bar{u}_{1,\varepsilon}, D^2\bar{u}_{1,\varepsilon}, \bar{u}_{1,\varepsilon}(\cdot)) \\ & \leq C(|\bar{u}_{1,\varepsilon}|_0 e^{-\ell R} + |D^3\bar{u}_{1,\varepsilon}|_0 r^{3-\alpha}) \\ & \leq C(e^{-\ell R} + \varepsilon^{-2} r^{3-\alpha}) =: C(r, R, \varepsilon), \end{aligned}$$

for all  $x \in \mathbb{R}^N$ . Since  $\bar{u}_{1,\varepsilon}$  is a subsolution of (6.2), we get

$$H(x, \bar{u}_{1,\varepsilon}(y), D\bar{u}_{1,\varepsilon}, D^2\bar{u}_{1,\varepsilon}, \bar{u}_{1,\varepsilon}(\cdot)) \leq C(r, R, \varepsilon) \quad \text{in } \mathbb{R}^N.$$

From this inequality and (A.3) we see that  $\bar{u}_{1,\varepsilon} - C(r, R, \varepsilon)/\lambda$  is a subsolution of (1.1), and by the comparison principle (Theorem 3.3) we have

$$\bar{u}_{1,\varepsilon} - u \leq C(r, R, \varepsilon)/\lambda \quad \text{in } \mathbb{R}^N.$$

In view of Theorem 7.2, (A.1), and (A.2), it is easy to see that  $|\bar{u}_1^\varepsilon - \bar{u}_1|_0 \leq C\varepsilon$ . So by properties of mollifiers we have

$$|\bar{u}_1 - \bar{u}_{1,\varepsilon}|_0 \leq C\varepsilon.$$

Thus we conclude that

$$u - \bar{u} = u - \bar{u}_{1,\varepsilon} + \bar{u}_{1,\varepsilon} - \bar{u} \geq -C(r, R, \varepsilon)/\lambda - C\varepsilon \quad \text{in } \mathbb{R}^N,$$

and the lower bound follows after a minimization over  $\varepsilon$ .  $\square$

## 7. APPENDIX: SOME TECHNICAL RESULTS.

In this appendix we state and prove some technical results that were used in the proofs of Theorems 4.4 and 6.2. To obtain results that encompass both (1.1) and (6.2), we consider an equation of the form

$$(7.1) \quad \sup_{\theta \in \Theta} \left\{ -\operatorname{tr} \left[ \left( a_1^\theta(x) + a_2(x) \right) D^2 u \right] - b^\theta(x) Du - \mathcal{J}_1^\theta u(x) + c^\theta(x)u - f^\theta(x) \right\} = 0,$$



where

$$\begin{aligned}
 \mathcal{J}_1^\theta u(x) &:= \int_E [u(x + \eta_1^\theta(x, z)) - u(x) - \mathbf{1}_{|z| < 1} \eta_1^\theta(x, z) Du(x)] \nu(dz), \\
 (7.2) \quad a_1^\theta(x) &:= \frac{1}{2} \sigma^\theta(x) \sigma^\theta(x)^\top, \\
 a_2^\theta(x) &:= \frac{1}{2} \int_{E_1} \eta_2^\theta(x, z) \eta_2^\theta(x, z)^\top \nu(dz).
 \end{aligned}$$

In this section we always assume that (A.1), (A.2), and (A.3') hold: we mean that the requirements on  $\eta$  in (A.2) are met by both  $\eta_1$  and  $\eta_2$ . We start by giving an existence, uniqueness and regularity result for this equation.

**Theorem 7.1.** *Assume (A.1), (A.2), and (A.3') hold.*

(i) *There exists a unique viscosity solution  $u \in C_b(\mathbb{R}^N)$  of equation (7.2) which is Hölder continuous, i.e., there is a  $\delta \in (0, 1]$  such that*

$$|u(x) - u(y)| \leq C|x - y|^\delta \quad \text{for all } x, y \in \mathbb{R}^N.$$

(ii) *There exists a constant  $\lambda_0$  depending only on  $N$ ,  $\sup_\theta [\sigma^\theta]_1$ ,  $\sup_\theta [b^\theta]_1$ ,  $\sup_\theta \int_E [\eta_1^\theta(\cdot, z)]_1^2 \nu(dz)$ , and  $\sup_\theta \int_E [\eta_2^\theta(\cdot, z)]_1^2 \nu(dz)$  such that if*

$$\lambda \geq \lambda_0,$$

*then the viscosity solution  $u$  of (7.2) is Lipschitz continuous ( $\delta = 1$  above).*

(iii) *Let  $u, -v \in USC_b(\mathbb{R}^N)$ . If  $u$  and  $v$  are respectively viscosity sub- and supersolutions of (7.2), then*

$$u \leq v \quad \text{in } \mathbb{R}^N.$$

We will not prove this result. The proof is similar to the proof of Theorem 3.3, the only difference is the term  $a_2^\theta$  which can be handled as shown in the proof of the next Theorem.

**Theorem 7.2.** *Let  $u, v \in C^{0,1}(\mathbb{R}^N)$  be viscosity solutions of (7.1) corresponding to the set of coefficients  $\{\sigma^\theta, c^\theta, b^\theta, f^\theta, \eta_1^\theta, \eta_2^\theta\}$  and  $\{\bar{\sigma}^\theta, \bar{c}^\theta, \bar{b}^\theta, \bar{f}^\theta, \bar{\eta}_1^\theta, \bar{\eta}_2^\theta\}$ , respectively. Suppose assumptions (A.1), (A.2), and (A.3') hold for both sets of coefficients with constants  $M$  and  $\bar{M}$ , respectively, and  $\bar{\lambda} = \lambda$ . Then there exists a constant  $\tilde{C} = \tilde{C}(M, \bar{M})$  such that the following continuous dependence estimate holds:*

$$\begin{aligned}
 &\lambda |u - v|_0 \\
 &\leq \tilde{C} \sup_{\theta \in \Theta} \left\{ |\bar{\sigma}^\theta - \sigma^\theta| + \left| \int_{E_1} (|\eta_1^\theta(\cdot, z) - \bar{\eta}_1^\theta(\cdot, z)|^2 + |\eta_2^\theta(\cdot, z) - \bar{\eta}_2^\theta(\cdot, z)|^2) \nu(dz) \right|_0^{1/2} \right\} \\
 &\quad + \sup_{\theta \in \Theta} \left\{ ([u]_1 \wedge [v]_1) \left[ \left| \int_{E_1} |\eta_1^\theta(\cdot, z) - \bar{\eta}_1^\theta(\cdot, z)| \nu(dz) \right|_0 + |\bar{b}^\theta - b^\theta|_0 \right] \right. \\
 &\quad \left. + (|u|_0 \wedge |v|_0) |\bar{c}^\theta - c^\theta|_0 + |\bar{f}^\theta - f^\theta|_0 \right\}.
 \end{aligned}$$

*Proof.* When  $\eta_1, \eta_2 \equiv 0$  (the pure diffusion case) this result is proved in [36], see also [37]. Otherwise, the proof is similar to the one given in [38]. To avoid too much repetition of [36, 37, 38], we assume that  $\sigma = \bar{\sigma}$  (and thus  $a_1 = \bar{a}_1$ ),  $b = \bar{b}$ ,  $c = \bar{c}$ , and  $f = \bar{f}$ , so that in the proof below we concentrate only on the integro part of the continuous dependence estimate.

We begin by introducing the functions

$$\phi(x, y) = \alpha|x - y|^2 + \varepsilon(|x|^2 + |y|^2), \quad \psi(x, y) = u(x) - v(y) - \phi(x, y).$$

Note that  $\psi$  is continuous and tends to  $-\infty$  as  $|x| + |y| \rightarrow \infty$ , so its global maximum  $M$  is attained at some point  $(\bar{x}, \bar{y}) \in \mathbb{R}^N \times \mathbb{R}^N$ . We will derive a positive upper bound for  $M$  in the case that  $M > 0$ .

By the defining viscosity subsolution/supersolution inequalities for  $u, v$  and the maximum principle for semicontinuous functions [23], properly adapted to integro-partial differential equations in [39], we get for any  $0 < \kappa < 1$  the inequality

$$\begin{aligned}
 & \lambda(u(\bar{x}) - v(\bar{y})) \\
 & \leq \sup_{\theta \in \Theta} \left\{ \frac{1}{2} \operatorname{tr} [a_1^\theta(\bar{x})X] - \frac{1}{2} \operatorname{tr} [a_1^\theta(\bar{y})Y] \right. \\
 & \quad + b^\theta(\bar{x})D_x \phi(\bar{x}, \bar{y}) - b^\theta(\bar{y})(-D_y \phi(\bar{x}, \bar{y})) \\
 & \quad + (c^\theta(\bar{y}) - c^\theta(\bar{x}))v(\bar{y}) + f^\theta(\bar{y}) - f^\theta(\bar{x}) \\
 & \quad + \frac{1}{2} \operatorname{tr} [a_2^\theta(\bar{x})X] - \frac{1}{2} \operatorname{tr} [\bar{a}_2^\theta(\bar{y})Y] \\
 & \quad + \mathcal{J}_\kappa^\theta \phi(\bar{x}, \bar{y}) - \bar{\mathcal{J}}_\kappa^\theta(-\phi)(\bar{x}, \bar{y}) \\
 & \quad \left. + \mathcal{J}^{\theta, \kappa}(u, \phi(\cdot, \bar{y}))(\bar{x}) - \bar{\mathcal{J}}^{\theta, \kappa}(v, -\phi(\bar{x}, \cdot))(\bar{y}) \right\},
 \end{aligned} \tag{7.3}$$

where  $\mathcal{J}_\kappa^\theta, \mathcal{J}^{\theta, \kappa}$  and  $\bar{\mathcal{J}}_\kappa^\theta, \bar{\mathcal{J}}^{\theta, \kappa}$  are defined as in (3.3) with  $\eta^\theta$  replaced by  $\eta_1^\theta$  and  $\bar{\eta}_1^\theta$  respectively, and  $X, Y$  are appropriate matrices given by the maximum principle for semicontinuous functions (see [39] for details). Here we have also used the estimate

$$\begin{aligned}
 c^\theta(\bar{x})u(\bar{x}) - c^\theta(\bar{y})v(\bar{y}) &= c^\theta(\bar{x})(u(\bar{x}) - v(\bar{y})) + (c^\theta(\bar{x}) - c^\theta(\bar{y}))v(\bar{y}) \\
 &\geq \lambda(u(\bar{x}) - v(\bar{y})) + (c^\theta(\bar{x}) - c^\theta(\bar{y}))v(\bar{y}),
 \end{aligned}$$

which follows from (A.3') and the fact that  $u(\bar{x}) - v(\bar{y}) \geq 0$  (since otherwise  $M \leq 0$ ).

A by now standard argument, see [36, 37, 38], using the Lipschitz regularity of  $u, v$  and the coefficients, yields that the first four differences on the right-hand side of (7.3) are bounded by

$$C\alpha^{-1} + \omega(\varepsilon),$$

for some modulus  $\omega$ . Similar arguments also lead to the estimate

$$\frac{1}{2} \operatorname{tr} [a_2^\theta(\bar{x})X] - \frac{1}{2} \operatorname{tr} [\bar{a}_2^\theta(\bar{y})Y] \leq \alpha \int_{E_1} |\eta_2(\bar{x}, z) - \bar{\eta}_2(\bar{y}, z)|^2 \nu(dz).$$

This follows from "Ishii's trick", see [23, Example 3.6].

By Lemma 3.1 the  $\mathcal{J}_\kappa^\theta - \bar{\mathcal{J}}_\kappa^\theta$  term is bounded by some modulus  $\omega_{\alpha, \varepsilon}(\kappa)$ . So we are only left with the  $\mathcal{J}^{\theta, \kappa} - \bar{\mathcal{J}}^{\theta, \kappa}$  term. We have

$$\begin{aligned}
 & \mathcal{J}^{\theta, \kappa}(u, \phi(\cdot, \bar{y}))(\bar{x}) - \bar{\mathcal{J}}^{\theta, \kappa}(v, -\phi(\bar{x}, \cdot))(\bar{y}) \\
 &= \int_{\kappa < |z| < 1} \left[ u(\bar{x} + \eta_1^\theta(\bar{x}, z)) - u(\bar{x}) - (v(\bar{y} + \bar{\eta}_1^\theta(\bar{y}, z)) - v(\bar{y})) \right. \\
 & \quad \left. - \eta_1^\theta(\bar{x}, z)D_x \phi(\bar{x}, \bar{y}) - \bar{\eta}_1^\theta(\bar{y}, z)D_y \phi(\bar{x}, \bar{y}) \right] \nu(dz) \\
 & \quad + \int_{|z| \geq 1} \left[ u(\bar{x} + \eta_1^\theta(\bar{x}, z)) - u(\bar{x}) - (v(\bar{y} + \bar{\eta}_1^\theta(\bar{y}, z)) - v(\bar{y})) \right] \nu(dz).
 \end{aligned} \tag{7.4}$$

Since

$$\psi(\bar{x}, \bar{y}) \geq \psi(\bar{x} + \eta_1^\theta(\bar{x}, z), \bar{y} + \bar{\eta}_1^\theta(\bar{y}, z)),$$

the first integral in (7.4) is bounded by

$$\int_{\kappa < |z| < 1} \left[ \alpha |\eta_1^\theta(\bar{x}, z) - \bar{\eta}_1^\theta(\bar{y}, z)|^2 + C\varepsilon(|z|^2 \mathbf{1}_{E_1} + e^{\ell|z|} \mathbf{1}_{E^1}) \right] \nu(dz),$$

where the last term follows from (A.2) and a Taylor expansion in  $x$  and  $y$  of the  $\varepsilon$ -terms. By the Lipschitz regularity of  $u, v$  and  $\eta_1$ , cfr. (A.2), and the integrability conditions satisfied by  $\nu(dz)$ , cfr. (A.1), we find that this term is bounded by

$$C\alpha \int_{E_1} |\eta_1^\theta(\bar{x}, z) - \bar{\eta}_1^\theta(\bar{x}, z)|^2 \nu(dz) + C\alpha^{-1} + C\varepsilon.$$

Consider now the integrand of the second term in (7.4). Assume that  $[u]_1 \leq [v]_1$ . By the Lipschitz regularity of  $u$  this integrand is bounded by

$$[u]_1 |\bar{x} + \eta_1^\theta(\bar{x}, z) - \bar{y} - \bar{\eta}_1^\theta(\bar{y}, z)| + (u(\bar{y} + \bar{\eta}_1^\theta(\bar{y}, z)) - v(\bar{y} + \bar{\eta}_1^\theta(\bar{y}, z))) - (u(\bar{x}) - v(\bar{y})).$$

Note that the second and third differences in this expression are upper bounded by

$$\sup_{\mathbb{R}^N} \{u - v\} - \sup_{\mathbb{R}^{2N}} \{u(x) - v(y) - \phi(x, y)\} - \phi(\bar{x}, \bar{y}),$$

By the Lipschitz regularity of  $u, v$  and the inequality  $\sup\{\dots\} - \sup\{\dots\} \leq \sup\{\dots - \dots\}$ , this in turn can be bounded by

$$\sup_{\mathbb{R}^{2N}} \{[u]_1 |x - y| - \alpha |x - y|^2\}.$$

We bound this expression by maximizing with respect to  $r = |x - y|$  and get  $C\alpha^{-1}$ . We can now conclude that

$$([u]_1 \wedge [v]_1 \text{Big}) \int_{E_1} |\eta_1^\theta(\bar{x}, z) - \bar{\eta}_1^\theta(\bar{x}, z)| \nu(dz) + C\alpha^{-1}$$

constitutes an upper bound on the second integral in (7.4). Here we also used (A.1), (A.2), and the standard bound  $|\bar{x} - \bar{y}| \leq C\alpha^{-1}$ .

Combining all the previous estimates we have achieved

$$\begin{aligned} \lambda M &\leq \lambda(u(\bar{x}) - v(\bar{y})) \\ &\leq C\alpha \sup_{\theta \in \Theta} \left| \int_{E_1} (|\eta_1^\theta(\cdot, z) - \bar{\eta}_1^\theta(\cdot, z)|^2 + |\eta_2^\theta(\cdot, z) - \bar{\eta}_2^\theta(\cdot, z)|^2) \nu(dz) \right|_0 \\ &\quad + ([u]_1 \wedge [v]_1) \sup_{\theta \in \Theta} \left| \int_{E_1} |\eta_1^\theta(\cdot, z) - \bar{\eta}_1^\theta(\cdot, z)| \nu(dz) \right|_0 + C\alpha^{-1} + \omega(\varepsilon) + \omega_{\alpha, \varepsilon}(\kappa). \end{aligned}$$

This last estimate holds trivially when  $M \leq 0$ . Since

$$\lambda|(u - v)^+|_0 \leq \lambda M,$$

the result of the theorem now follows by first sending  $\kappa \rightarrow 0$ , then sending  $\varepsilon \rightarrow 0$ , then minimizing the resulting expression with respect to  $\alpha$ , and finally interchanging the roles of  $u$  and  $v$ .  $\square$

To prove the next theorem, we need the following auxilliary lemma:

**Lemma 7.3.** *Suppose (A.1), (A.2) hold and  $\lambda_1, \dots, \lambda_n \geq 0$  satisfy  $\sum_{i=1}^n \lambda_i = 1$ . If  $u^1, \dots, u^n \in C_b(\mathbb{R}^N)$  are viscosity subsolutions of (7.1), then  $\sum_{i=1}^n \lambda_i u^i$  is also a viscosity subsolution of (7.1).*

*Proof.* In this proof we set  $a := a_1 + a_2$ . We prove the result for  $n = 2$ . The general case follows by induction. We will show that  $\lambda_1 u^1 + \lambda_2 u^2$  is a subsolution of (7.1). Let  $\chi \in C^2(\mathbb{R}^N)$  and assume that  $\lambda_1 u^1 + \lambda_2 u^2 - \chi$  has a global maximum at  $\bar{x} \in \mathbb{R}^N$ . We may assume without loss of generality that this maximum is strict. Now introduce the functions

$$\begin{aligned} \phi(x, y) &= \frac{1}{\alpha} |x - y|^2, \\ \psi(x, y) &= \lambda_1 u^1(x) + \lambda_2 u^2(y) - \lambda_1 \chi(x) - \lambda_2 \chi(y) - \phi(x, y). \end{aligned}$$

Let  $m_\alpha = \sup_{\mathbb{R}^{2N}} \psi(x, y)$  and note that this supremum is attained at some point  $(x_\alpha, y_\alpha)$  when  $\alpha$  is small. Moreover, it is possible to show that  $x_\alpha, y_\alpha \rightarrow \bar{x}$  and  $\frac{1}{\alpha}|x_\alpha - y_\alpha|^2 \rightarrow 0$  as  $\alpha \rightarrow 0$ .

Now by the maximum principle for semicontinuous functions [23], properly adapted to integro-partial differential equations in [39],

$$(7.5) \quad \sup_{\theta \in \Theta} \left\{ -\frac{1}{2} \text{tr} \left[ a^\theta(x_\alpha)X + a^\theta(y_\alpha)Y \right] \right. \\
\begin{aligned}
& - b^\theta(x_\alpha)(\lambda_1 D\chi(x_\alpha) + D_x \phi(x_\alpha, y_\alpha)) \\
& - b^\theta(y_\alpha)(\lambda_2 D\chi(y_\alpha) + D_y \phi(x_\alpha, y_\alpha)) \\
& + c^\theta(x_\alpha)\lambda_1 u^1(x_\alpha) + c^\theta(y_\alpha)\lambda_2 u^2(y_\alpha) \\
& - \mathcal{J}_\kappa^\theta(\lambda_1 \chi + \phi(\cdot, y_\alpha))(x_\alpha) - \mathcal{J}_\kappa^\theta(\lambda_2 \chi + \phi(x_\alpha, \cdot))(y_\alpha) \\
& - \mathcal{J}^{\theta, \kappa}(\lambda_1 u^1, \lambda_1 D\chi(x_\alpha) + D_x \phi(\cdot, y_\alpha))(x_\alpha) \\
& - \mathcal{J}^{\theta, \kappa}(\lambda_2 u^2, \lambda_2 D\chi(y_\alpha) + D_y \phi(x_\alpha, \cdot))(y_\alpha) \\
& \left. - \lambda_1 f^\theta(x_\alpha) - \lambda_2 f^\theta(y_\alpha) \right\} \leq 0,
\end{aligned}$$

for some matrices  $X, Y$  satisfying

$$\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \leq \frac{C}{\alpha} \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + \begin{pmatrix} \lambda_1 D^2 \chi(x_\alpha) & 0 \\ 0 & \lambda_2 D^2 \chi(y_\alpha) \end{pmatrix}.$$

We immediately see that

$$\text{tr} \left[ a^\theta(x_\alpha)X + a^\theta(y_\alpha)Y \right] \leq \text{tr} \left[ a^\theta(x_\alpha)\lambda_1 D^2 \chi(x_\alpha) + a^\theta(y_\alpha)\lambda_2 D^2 \chi(y_\alpha) \right].$$

Next observe that  $\psi(\bar{x}, \bar{y}) \geq \psi(\bar{x} + \eta^\theta(\bar{x}, z), \bar{y} + \eta^\theta(\bar{y}, z))$  which is equivalent to

$$\begin{aligned}
& \lambda_1 u^1(\bar{x} + \eta^\theta(\bar{x}, z)) - \lambda_1 u^1(\bar{x}) + \lambda_2 u^2(\bar{y} + \eta^\theta(\bar{y}, z)) - \lambda_2 u^2(\bar{y}) \\
& \leq \lambda_1 \chi(\bar{x} + \eta^\theta(\bar{x}, z)) - \lambda_1 \chi(\bar{x}) + \lambda_2 \chi(\bar{y} + \eta^\theta(\bar{y}, z)) - \lambda_2 \chi(\bar{y}) \\
& + \text{terms involving } \phi,
\end{aligned}$$

and hence

$$(7.6) \quad \begin{aligned}
& \mathcal{J}^{\theta, \kappa}(\lambda_1 u^1, \lambda_1 D\chi(x_\alpha))(x_\alpha) + \mathcal{J}^{\theta, \kappa}(\lambda_2 u^2, \lambda_2 D\chi(y_\alpha))(y_\alpha) \\
& \leq \mathcal{J}^{\theta, \kappa}(\lambda_1 \chi, \lambda_1 D\chi(x_\alpha))(x_\alpha) + \mathcal{J}^{\theta, \kappa}(\lambda_2 \chi, \lambda_2 D\chi(y_\alpha))(y_\alpha) \\
& + \text{terms involving } \phi.
\end{aligned}$$

Now computations similar to those in the proof of Theorem 7.2 show that the sum of all terms containing  $\phi$  in (7.5) and (7.6) can be bounded by

$$\frac{C}{\alpha} |x_\alpha - y_\alpha|^2 + \omega_\alpha(\kappa).$$

Hence we are left with the following inequality:

$$\begin{aligned}
& \sup_{\theta \in \Theta} \left\{ \text{tr} \left[ a^\theta(x_\alpha)\lambda_1 D^2 \chi(x_\alpha) + a^\theta(y_\alpha)\lambda_2 D^2 \chi(y_\alpha) \right] \right. \\
& - b^\theta(x_\alpha)\lambda_1 D\chi(x_\alpha) - b^\theta(y_\alpha)\lambda_2 D\chi(y_\alpha) \\
& + c^\theta(x_\alpha)\lambda_1 u^1(x_\alpha) + c^\theta(y_\alpha)\lambda_2 u^2(y_\alpha) \\
& - \lambda_1 \mathcal{J}^{\theta, \kappa}(\chi, D\chi(x_\alpha))(x_\alpha) - \lambda_2 \mathcal{J}^{\theta, \kappa}(\chi, D\chi(y_\alpha))(y_\alpha) \\
& \left. - \lambda_1 f^\theta(x_\alpha) - \lambda_2 f^\theta(y_\alpha) \right\} \leq \frac{C}{\alpha} |x_\alpha - y_\alpha|^2 + \omega_\alpha(\kappa).
\end{aligned}$$

Sending first  $\kappa \rightarrow 0$  and then  $\alpha \rightarrow 0$  we get

$$\sup_{\theta \in \Theta} \left\{ -\frac{1}{2} \operatorname{tr} \left[ a^\theta(\bar{x}) D^2 \chi(\bar{x}) \right] - b^\theta(\bar{x}) D\chi(\bar{x}) \right. \\ \left. + c^\theta(\bar{x}) \left( \lambda_1 u^1(\bar{x}) + \lambda_2 u^2(\bar{x}) \right) - \mathcal{J}^{\theta,0}(\chi, D\chi(\bar{x}))(\bar{x}) \right\} \leq 0.$$

Since  $\mathcal{J}^{\theta,0}(\chi, D\chi(\bar{x}))(\bar{x}) = \mathcal{J}^\theta \chi(\bar{x})$  (see (1.3)), according to Definition 3.1,  $\lambda_1 u^1 + \lambda_2 u^2$  is a viscosity subsolution of (7.1).  $\square$

Before we can state the final result of this section, we need to introduce an equation which has both (1.5) and the equation appearing in the proof of Theorem 6.2 as special cases. To this end, we introduce the equation

$$(7.7) \quad \sup_{\substack{\theta \in \Theta \\ |e| \leq \varepsilon}} \left\{ -\operatorname{tr} \left[ \left( a_1^\theta(x+e) + a_2(x+e) \right) D^2 u \right] - b^\theta(x+e) Du - \mathcal{J}_1^{\theta,e} u(x) \right. \\ \left. + c^\theta(x+e)u - f^\theta(x+e) \right\} = 0 \quad \text{in } \mathbb{R}^N,$$

where  $a_1, a_2$  are defined in (7.2) and

$$\mathcal{J}_1^{\theta,e} u(x) := \int_E \left[ u(x + \eta_1^\theta(x+e, z)) - u(x) - \mathbf{1}_{\{|z| < 1\}} \eta_1^\theta(x+e, z) Du(x) \right] \nu(dz).$$

The final result of this section is

**Theorem 7.4.** *Suppose (A.1) - (A.3) hold and let  $u^\varepsilon$  be a  $C^{0,1}(\mathbb{R}^N)$  viscosity subsolution of (7.7). Then*

$$u_\varepsilon := \rho_\varepsilon * u^\varepsilon$$

*is a viscosity subsolution of (7.1).*

*Proof.* We start by observing that from equation (7.7) we have

$$\sup_{\theta \in \Theta} \left\{ -\operatorname{tr} \left[ \left( a_1^\theta(x+e) + a_2(x+e) \right) D^2 u^\varepsilon \right] - b^\theta(x+e) Du^\varepsilon - \mathcal{J}_1^{\theta,e} u^\varepsilon(x) \right. \\ \left. + c^\theta(x+e)u^\varepsilon - f^\theta(x+e) \right\} = 0 \quad \text{in } \mathbb{R}^N \quad \text{for every } |e| \leq \varepsilon.$$

By a change of variables  $x \rightarrow x - e$  we see that  $u^\varepsilon(\cdot - e)$  is a viscosity subsolution of (7.1) for every  $|e| \leq \varepsilon$ . Now we define the following quantities:

$$Q_\alpha(h) := \alpha + \left[ -\frac{h}{2}, \frac{h}{2} \right]^N, \\ \rho_{\varepsilon,h}(\alpha) := \int_{Q_\alpha(h)} \rho_\varepsilon(y) dy, \\ u_{\varepsilon,h}(x) := \sum_{\alpha \in h\mathbb{Z}^N} u^\varepsilon(x - \alpha) \rho_{\varepsilon,h}(\alpha).$$

The function  $u_{\varepsilon,h}$  is an approximation of  $u_\varepsilon$  and by properties of mollifiers we see that  $u_{\varepsilon,h} \rightarrow u_\varepsilon$  uniformly as  $h \rightarrow 0$ . Note that  $\rho_{\varepsilon,h}(\alpha) \geq 0$  and that  $\rho_{\varepsilon,h}(\alpha) > 0$  only for finitely many  $\alpha$  since  $\rho_\varepsilon$  has compact support. Furthermore

$$\sum_{\alpha \in h\mathbb{Z}^N} \rho_{\varepsilon,h}(\alpha) = \int_{\mathbb{R}^N} \rho_\varepsilon(y) dy = 1,$$

and therefore when  $h < \varepsilon$ ,  $u_{\varepsilon,h}(x)$  is a convex combination of viscosity subsolutions  $u^\varepsilon(\cdot - \alpha)$  of (7.1). Thanks to Lemma 7.3, we conclude that  $u_{\varepsilon,h}$  itself is a viscosity subsolution of (7.1). Hence, by the stability result for viscosity subsolutions, cf. [23, Lemma 6.1],  $u_\varepsilon = \lim_{h \rightarrow 0} u_{\varepsilon,h}$  is also a viscosity subsolution of (7.1).  $\square$

## REFERENCES

- [1] O. Alvarez and A. Tourin. Viscosity solutions of nonlinear integro-differential equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 13(3):293–317, 1996.
- [2] A. L. Amadori. The obstacle problem for nonlinear integro-differential operators arising in option pricing. Quaderno IAC Q21-000, 2000.
- [3] A. L. Amadori. Nonlinear integro-differential evolution problems arising in option pricing: a viscosity solutions approach. *Differential Integral Equations*, 16(7):787–811, 2003.
- [4] A. L. Amadori, K. H. Karlsen, and C. La Chioma. Non-linear degenerate integro-partial differential evolution equations related to geometric Lévy processes and applications to backward stochastic differential equations. *Stoch. Stoch. Rep.*, 76(2):147–177, 2004.
- [5] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.
- [6] G. Barles. *Solutions de viscosité des équations de Hamilton-Jacobi*, volume 17 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Paris, 1994.
- [7] G. Barles, R. Buckdahn, and E. Pardoux. Backward stochastic differential equations and integral-partial differential equations. *Stochastics Stochastics Rep.*, 60(1-2):57–83, 1997.
- [8] G. Barles and E. R. Jakobsen. On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations. *M2AN Math. Model. Numer. Anal.*, 36(1):33–54, 2002.
- [9] G. Barles and E. R. Jakobsen. Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.* To appear.
- [10] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4(3):271–283, 1991.
- [11] O. E. Barndorff-Nielsen. Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society London*, 353:401–419, 1977.
- [12] O. E. Barndorff-Nielsen. Processes of normal inverse Gaussian type. *Finance Stoch.*, 2(1):41–68, 1998.
- [13] F. E. Benth, K. H. Karlsen, and K. Reikvam. A note on portfolio management under non-Gaussian logreturns. *Int. J. Theor. Appl. Finance*, 4(5):711–731, 2001.
- [14] F. E. Benth, K. H. Karlsen, and K. Reikvam. Optimal portfolio management rules in a non-Gaussian market with durability and intertemporal substitution. *Finance Stoch.*, 5(4):447–467, 2001.
- [15] F. E. Benth, K. H. Karlsen, and K. Reikvam. Optimal portfolio selection with consumption and nonlinear integro-differential equations with gradient constraint: a viscosity solution approach. *Finance Stoch.*, 5(3):275–303, 2001.
- [16] F. E. Benth, K. H. Karlsen, and K. Reikvam. Portfolio optimization in a Lévy market with intertemporal substitution and transaction costs. *Stoch. Stoch. Rep.*, 74(3-4):517–569, 2002.
- [17] J. F. Bonnans and H. Zidani. Consistency of generalized finite difference schemes for the stochastic Bellman equation. *SIAM J. Numer. Anal.*, 41(3):1008–1021 (electronic), 2003.
- [18] M. Briani, C. La Chioma, and R. Natalini. Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in finance theory. *Numer. Math.* 98(4):607–646, 2004.
- [19] M. Briani, R. Natalini, and G. Russo. Implicit-explicit numerical schemes for jump-diffusion processes. Technical report, IAC-CNR, Roma, 2004.
- [20] F. Camilli and M. Falcone. An approximation scheme for the optimal control of diffusion processes. *RAIRO Modél. Math. Anal. Numér.*, 29(1):97–122, 1995.
- [21] R. Cont and P. Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.
- [22] R. Cont and E. Voltchkova. Finite difference methods for option pricing in jump-diffusion and exponential lévy models. *SIAM J. Numer. Anal.* To appear.
- [23] M. G. Crandall, H. Ishii, and P.-L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.
- [24] M. G. Crandall and P.-L. Lions. Two approximations of solutions of Hamilton-Jacobi equations. *Math. Comp.*, 43(167):1–19, 1984.
- [25] P. J. Davis and P. Rabinowitz. *Methods of numerical integration*. Computer Science and Applied Mathematics. Academic Press Inc., Orlando, FL, 1984.
- [26] Y. d’Halluin, P. A. Forsyth, and G. Labahn. A penalty method for American options with jump-diffusion processes. *Numer. Math.*, 97(2):321–352, 2004.
- [27] P. Dupuis and M. R. James. Rates of convergence for approximation schemes in optimal control. *SIAM J. Control Optim.*, 36(2):719–741 (electronic), 1998.

- [28] E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 1:281–299, 1995.
- [29] S. Elganjoui. Master’s thesis, Department of Mathematics, University of Bergen, January 2001.
- [30] S. Elganjoui and K. H. Karlsen. A markov chain approximation scheme for a singular investment-consumption problem with lévy driven stock prices. available at the URL <http://www.mi.uib.no/~kennethk/>, 2002.
- [31] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*. Springer-Verlag, New York, 1993.
- [32] M. G. Garroni and J.-L. Menaldi. *Green functions for second order parabolic integro-differential problems*, volume 275 of *Pitman Research Notes in Mathematics Series*. Longman Scientific & Technical, Harlow, 1992.
- [33] M. G. Garroni and J. L. Menaldi. *Second order elliptic integro-differential problems*, volume 430 of *Chapman & Hall/CRC Research Notes in Mathematics*. Chapman & Hall/CRC, Boca Raton, FL, 2002.
- [34] H. Ishii and P.-L. Lions. Viscosity solutions of fully nonlinear second-order elliptic partial differential equations. *J. Differential Equations*, 83(1):26–78, 1990.
- [35] E. R. Jakobsen. On the rate of convergence of approximation schemes for Bellman equations associated with optimal stopping time problems. *Math. Models Methods Appl. Sci.*, 13(5):613–644, 2003.
- [36] E. R. Jakobsen and K. H. Karlsen. Continuous dependence estimates for viscosity solutions of fully nonlinear degenerate elliptic equations. *Electron. J. Differential Equations*, pages No. 39, 10 pp. (electronic), 2002.
- [37] E. R. Jakobsen and K. H. Karlsen. Continuous dependence estimates for viscosity solutions of fully nonlinear degenerate parabolic equations. *J. Differential Equations*, 183(2):497–525, 2002.
- [38] E. R. Jakobsen and K. H. Karlsen. Continuous dependence estimates for viscosity solutions of integro-pdes. *J. Differential Equations* 212(2): 278–318, 2005.
- [39] E. R. Jakobsen and K. H. Karlsen. A “maximum principle for semicontinuous functions” applicable to integro-partial differential equations. *NoDEA Nonlinear Differential Equations Appl.* To appear.
- [40] N. V. Krylov. On the rate of convergence of finite-difference approximations for Bellman’s equations. *Algebra i Analiz*, 9(3):245–256, 1997.
- [41] N. V. Krylov. On the rate of convergence of finite-difference approximations for Bellman’s equations with variable coefficients. *Probab. Theory Related Fields*, 117(1):1–16, 2000.
- [42] N. V. Krylov. On the rate of convergence of finite-difference approximations for bellman equations with lipschitz coefficients. Preprint, 2004.
- [43] H. J. Kushner and P. G. Dupuis. *Numerical methods for stochastic control problems in continuous time*. Springer-Verlag, New York, 1992.
- [44] P.-L. Lions and B. Mercier. Approximation numérique des équations de Hamilton-Jacobi-Bellman. *RAIRO Anal. Numér.*, 14(4):369–393, 1980.
- [45] A.-M. Matache, T. von Petersdorff, and C. Schwab. Fast deterministic pricing of options on Lévy driven assets. *M2AN Math. Model. Numer. Anal.*, 38(1):37–71, 2004.
- [46] J.-L. Menaldi. Some estimates for finite difference approximations. *SIAM J. Control Optim.*, 27(3):579–607, 1989.
- [47] R. Mikulevičius and G. Pragarauskas. Nonlinear potentials of the Cauchy-Dirichlet problem for the Bellman integro-differential equation. *Liet. Mat. Rink.*, 36(2):178–218, 1996.
- [48] R. Mikulevičius and H. Pragarauskas. On the existence of a viscosity solution to the integro-differential Bellman equation. In *Probability theory and mathematical statistics (Vilnius, 1993)*, pages 553–564. TEV, Vilnius, 1994.
- [49] R. Mikulyavichyus and G. Pragarauskas. Nonlinear potentials of the Cauchy-Dirichlet problem for the Bellman integro-differential equation. *Liet. Mat. Rink.*, 36(2):178–218, 1996.
- [50] B. Øksendal and A. Sulem. Optimal consumption and portfolio with both fixed and proportional transaction costs. *SIAM J. Control Optim.*, 40(6):1765–1790 (electronic), 2002.
- [51] H. Pham. Optimal stopping of controlled jump-diffusion processes: a viscosity solution approach. *J. Math. Systems Estim. Control*, 8(1):27 pp. (electronic), 1998.
- [52] T. H. Rydberg. The normal inverse Gaussian Lévy process: simulation and approximation. *Comm. Statist. Stochastic Models*, 13(4):887–910, 1997.
- [53] A. Sayah. Équations d’Hamilton-Jacobi du premier ordre avec termes intégral-différentiels. I. Unicité des solutions de viscosité. *Comm. Partial Differential Equations*, 16(6-7):1057–1074, 1991.
- [54] A. Sayah. Équations d’Hamilton-Jacobi du premier ordre avec termes intégral-différentiels. II. Existence de solutions de viscosité. *Comm. Partial Differential Equations*, 16(6-7):1075–1093, 1991.

- [55] A. N. Shiryaev. *Essentials of stochastic finance*, volume 3 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Inc., River Edge, NJ, 1999. Facts, models, theory, Translated From the Russian man UScript by N. Kruzhilin.
- [56] H. M. Soner. Optimal control with state-space constraint. II. *SIAM J. Control Optim.*, 24(6):1110–1122, 1986.
- [57] H. M. Soner. Optimal control of jump-Markov processes and viscosity solutions. In *Stochastic differential systems, stochastic control theory and applications (Minneapolis, Minn., 1986)*, volume 10 of *IMA Vol. Math. Appl.*, pages 501–511. Springer, New York, 1988.

(Espen R. Jakobsen), DEPARTMENT OF MATHEMATICAL SCIENCES, NORWEGIAN UNIVERSITY OF SCIENCE TECHNOLOGY, N-7491 TRONDHEIM, NORWAY

*E-mail address:* `erj@math.ntnu.no`

*URL:* `http://www.math.ntnu.no/~erj/`

(Kenneth H. Karlsen), CENTRE OF MATHEMATICS FOR APPLICATIONS (CMA), DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OSLO, P.O. BOX 1053, BLINDERN, N-0316 OSLO, NORWAY

*E-mail address:* `kennethk@math.uio.no`

*URL:* `http://www.math.uio.no/~kennethk/`

(Claudia La Chioma), CAPITALIA - LINEA FINANZA, PRODOTTI E CANALI, QUANTITATIVE FINANCE, VIA G. PAISIELLO, 5, I-00198 ROMA, ITALIA

*E-mail address:* `claudia.lachioma@capitalia.it`

*URL:* `http://www.math.uio.no/~claudlc/`